

Aalto-yliopisto

Insinööritieteiden korkeakoulu

Liikennetekniikan koulutusohjelma

Matti-Pekka Laaksonen

Liittymien liikenneturvallisuuden mittaaminen:

Vaarallisten kohteiden tunnistaminen Helsingin kantakaupungissa

Diplomityö

Espoo, 13. huhtikuuta 2015

Valvoja: Professori Tapio Luttinen, Aalto-yliopisto

Ohjaaja: DI Hanna Strömmer, Helsingin kaupunki

Aalto-yliopisto

Insinööritieteiden korkeakoulu

Liikennetekniikan koulutusohjelma

DIPLOMITYÖN

TIIVISTELMÄ

Tekijä:	Matti-Pekka Laaksonen		
Työn nimi:	Liittymien liikenneturvallisuuden mittaaminen: Vaarallisten kohteiden tunnistaminen Helsingin kantakaupungissa		
Päiväys:	13. huhtikuuta 2015	Sivumäärä:	81
Pääaine:	Liikennetekniikka	Koodi:	Yhd-71
Valvoja:	Professori Tapio Luttinen, Aalto-yliopisto		
Ohjaaja:	DI Hanna Strömmer, Helsingin kaupunki		
<p>Tutkimuksessa on sovellettu uutta menetelmää liikenneturvallisuuden mittaamiseksi ja mustien pisteiden, eli vaarallisten kohteiden tunnistamiseksi Helsingin kantakaupungin katuliittymissä. Työssä on aiemman tutkimuksen avulla verrattu eri menetelmiä mustien pisteiden tunnistamiseksi ja perusteltu käytetyn Empirical Bayes -menetelmän valinta.</p> <p>Tutkimuksessa sovellettava Empirical Bayes -menetelmä mittaa liikenneturvallisuutta liittymässä havaitun onnettomuusfrekvenssin sekä muilta vastaavilta liittymiltä odotetun onnettomuusfrekvenssin painotettuna keskiarvona. Vastaavien kohteiden onnettomuusfrekvenssi määritetään työssä laadittavalla tilastollisella mallilla.</p> <p>Tutkimusaineistona on käytetty poliisin tietoon tulleita onnettomuuksia sekä Helsingin kaupunkisuunnitteluviraston ylläpitämää liittymäpisterekisteriä. Mallintamisessa käytetyt muuttujat on luotu yhdistelemällä valmiiksi saatavilla olevaa tietoa liittymien ominaisuuksista.</p> <p>Työssä on laadittu yleistetty lineaarinen malli onnettomuusfrekvenssin ennustamiseksi. Laaditun negatiivisen binomiregressiomallin selittävinä muuttujina ovat liittymän pää- ja sivusuuntien liikennemäärät sekä liittymän haarojen lukumäärä. Liikennevalo-ohjauksisille ja valo-ohjauksettomille liittymille on laadittu erilliset regressiomallit.</p>			
Asiasanat:	liikenneturvallisuus, musta piste, Empirical Bayes, yleistetty lineaarinen malli, negatiivinen binomimalli		
Kieli:	Suomi		

Aalto University

School of Engineering

Degree Programme in Transportation Engineering

ABSTRACT OF
MASTER'S THESIS

Author:	Matti-Pekka Laaksonen		
Title:	Measuring the safety performance of intersections: Identification of accident black spots in the inner city of Helsinki		
Date:	April 13, 2015	Pages:	81
Major:	Transportation Engineering	Code:	Yhd-71
Supervisor:	Professor Tapio Luttinen, Aalto University		
Advisor:	M.Sc. (Tech.) Hanna Strömmer, City of Helsinki		
<p>A new method for the identification of accident hot spots in the intersections of Helsinki is applied in this master’s thesis. Prior research is used to compare different identification methods, and a method called Empirical Bayes is chosen for the analysis.</p> <p>The Empirical Bayes method estimates the safety of an entity by combining prior information from other sites to observed information from the observed entity. The prior information is obtained through the use of a statistical model.</p> <p>The negative binomial regression model used in the research is estimated using empirical accident data and detailed information about the intersections in Helsinki. The scope of research is limited to intersections in the inner city of Helsinki.</p> <p>The independent variables in the model are the AADTs of the intersection’s main and secondary legs and the number of legs in the intersection. Separate models are used for signalized and non-signalized intersections.</p>			
Keywords:	traffic safety, black spot, Empirical Bayes, generalized linear model, negative binomial		
Language:	Finnish		

Esipuhe

Tämä diplomityö on laadittu Helsingin kaupunkisuunnitteluviraston liikennejärjestelmätoimiston tilaamana. Työn ohjaajana on toiminut liikenneinsinööri, DI Hanna Strömmer Helsingin kaupunkisuunnitteluvirastosta ja työn valvojana professori Tapio Luttinen Aalto-yliopistosta. Työtä on ohjannut myös ohjausryhmä, johon ovat kuuluneet valvojan ja ohjaajan lisäksi kaupunkisuunnitteluviraston liikennejärjestelmätoimiston johtaja Matti Kivelä sekä liikennemalleista vastannut liikenneinsinööri Petri Blomqvist.

Haluan kiittää työn valvojaa professori Tapio Luttista. Haluan erityisesti kiittää työn ohjaajaa Hanna Strömmeriä jatkuvasta kannustuksesta, tuesta ja kärsivällisyydestä. Kiitän myös Matti Kivelää ohjausryhmän tapaamisista sekä liikenneinsinööri Hannu Seppälää työssä käytetyn liikennemääräaineiston muodostamisesta.

Espoo, 13. huhtikuuta 2015

Matti-Pekka Laaksonen

Sisällysluettelo

1	Johdanto	1
1.1	Tutkimuksen tausta	1
1.2	Tutkimuksen tavoite	2
1.3	Työn rajaukset	3
1.4	Liikenneturvallisuuden määrittäminen	4
2	Liikenneturvallisuuden mittaaminen	7
2.1	Mustien pisteiden tunnistaminen	7
2.1.1	Mustan pisteen määrittely	7
2.1.2	Satunnaisvaihtelun huomioiminen	8
2.2	Aiempi tutkimus aiheesta	9
2.2.1	Verrattavien menetelmien esittely	9
2.2.2	Merkittävimmät tutkimukset	10
2.2.3	Johtopäätökset kirjallisuudesta	13
2.2.4	Empirical Bayes	13
2.3	Onnettomuuksien satunnaisuus	16
2.3.1	Onnettomuuksien todennäköisyysjakauma	16

2.3.2	Ylihajonta	17
3	Tutkimusmenetelmät	19
3.1	Mustien pisteiden tunnistaminen tutkimuksessa	19
3.2	Onnettomuuksien mallintaminen	20
3.2.1	Yleistetty lineaarinen malli	20
3.2.2	Mallin jakaumaoletus	21
3.2.3	Mallin rakenteen valinta	21
3.2.4	Vastemuuttujan valinta	23
3.3	Mallien arviointi	25
3.3.1	Multikollineaarisuus	25
3.3.2	Mallien tilastollinen merkitsevyys	26
3.4	Tulosten arviointi	27
4	Tutkimusaineisto	29
4.1	Käytetyt aineistot	29
4.1.1	Onnettomuusaineisto	29
4.1.2	Liittymäpiste- ja liikennemääräaineisto	30
4.2	Aineiston rajausta	31
4.2.1	Käytettävän aineiston valinta	31
4.2.2	Alueellinen ja ajallinen rajausta	31
4.3	Tutkimusaineiston käsittely	33
4.3.1	Aineistosta luodut muuttujat	33
4.3.2	Hylätyt muuttujat	35
4.4	Muuttujien arviointi	37

4.4.1	Muuttujien korrelaatiot	37
5	Tutkimusprosessi ja tutkimuksen tulokset	39
5.1	Onnettomuusmallien laatiminen	39
5.1.1	Mallintamisprosessin kuvaus	39
5.1.2	Liikennemäärämuuttujan arviointi	42
5.2	Valitut onnettomuusmallit	45
5.2.1	Mallien esittely	45
5.2.2	Mallien arviointi	46
5.2.3	Regressiokertoimien tulkinta	47
5.3	Mustat pisteet	48
5.3.1	Empirical Bayes -menetelmällä tunnistetut kohteet	48
5.3.2	Empirical Bayes -menetelmän arviointi	50
5.4	IND5- ja EB-menetelmien vertailu	51
6	Yhteenveto ja johtopäätökset	53
6.1	Yhteenveto	53
6.2	Johtopäätökset	55
6.3	Ehdotukset jatkotoimenpiteiksi	56
6.3.1	Menetelmän laajentaminen muualle Helsinkiin	56
6.3.2	Menetelmän laajentaminen eri liittymätyyppeihin	56
6.3.3	Liittymäkohtaisten tietojen tarkentaminen	57
A	Onnettomuusmallit	63

Luku 1

Johdanto

1.1 Tutkimuksen tausta

Liikenneturvallisuuden parantamiseksi on kyettävä tunnistamaan sellaiset kohteet joissa liikenneturvallisuus on poikkeuksellisen huono. Näistä kohteista käytetään usein nimitystä musta piste, engl. *black spot*. Mustien pisteiden luotettavalla tunnistamisella voidaan saavuttaa parempi kokonaiskuva liikenneturvallisuuden tilasta sekä kohdentaa turvallisuutta parantavat toimenpiteet tehokkaasti. Tienpidon resurssien pienentyessä on tärkeää, että investoinnit liikenneturvallisuuteen tuottavat mahdollisimman suuren hyödyn.

Nykyisin Helsingissä määritellään liittymien liikenneturvallisuuden taso IND5-indeksiluvun avulla. Luku lasketaan jokaisen liittymän vuosittain tapahtuneista liikenneonnettomuuksista huomioimalla havaitut henkilövahinko-onnettomuudet painokertoimella 1 ja pelkkiin omaisuusvahinkoihin johtaneet onnettomuudet painokertoimella 0,2.

Satunnaisvaihtelun tasoittamiseksi IND5-luku lasketaan kolmen vuoden liukuvana keskiarvona. Menetelmän tunnistamille, mustiksi pisteiksi luokiteltaville liittymille lasketaan lisäksi IND5-aste, jossa indeksiluku on suhteutettu altistukseen, eli liittymien vuosittaiseen ajoneuvoliikenteen määrään.

IND5-luku on kuitenkin varsinainen luokitusmenetelmä, jolla vaarallisimmat kohteet tunnistetaan Helsingissä.

Sekä akateemisessa tutkimuksessa että muiden maiden viranomaistyössä on kuitenkin tunnistettu muita, tehokkaampia menetelmiä mustien pisteiden tunnistamiseksi. Tämän tutkimuksen aiheena on uuden mustien pisteiden tunnistamismenetelmän soveltaminen Helsingin katuliittymien liikenneturvallisuuden arvioinnissa.

1.2 Tutkimuksen tavoite

Työn teoreettisen osan tavoitteena on vertailla liikenneturvallisuuden mittamiseksi ja mustien pisteiden tunnistamiseksi käytettyjä menetelmiä aiemman tutkimuksen pohjalta. Vertailulla perustellaan työssä käytettävän Empirical Bayes -menetelmän valinta.

Työn empiirisen osan tavoitteena on soveltaa valittua mustien pisteiden tunnistamismenetelmää Helsingin kantakaupungin liittymiin. Työssä laaditaan menetelmän edellyttämä tilastollinen malli, jolla ennustetaan onnettomuuksien lukumäärää liittymissä.

Työssä pyritään vastaamaan seuraaviin tutkimuskysymyksiin:

1. Onko Empirical Bayes -menetelmä sovellettavissa Helsingin liittymien liikenneturvallisuuden arviointiin?
2. Millä tavoin liikenneonnettomuuksien lukumäärää Helsingin liittymissä voidaan mallintaa tilastollisesti?
3. Onko aiemmin käytetyn ja työssä tarkasteltavan menetelmän välillä havaittavissa eroa mustien pisteiden tunnistamisessa tarkasteltaessa Helsingin liittymiä?

Työn lopullisena päämääränä on osoittaa aiempaa kehittyneemmän mustien

pisteiden tunnistamismenetelmän soveltuvuus Helsingin liikenneturvallisuu-
den arviointiin.

1.3 Työn rajaukset

Tässä diplomityössä tarkastellaan vain liittymien liikenneturvallisuutta. Helsingissä kaksi kolmannesta kaikista liikenneonnettomuuksista tapahtuu liittymissä, joten liittymien vaikutus koko kaupungin liikenneturvallisuuteen on merkittävä (Helsingin kaupunki, 2012). Myös nykyinen IND5-menetelmä tarkastelee vain liittymien liikenneturvallisuutta, joten raja-
aus on perusteltu myös yhdenmukaisuuden vuoksi. Liittymien väliset katuosuudet sekä muut kohteet, kuten pysäköintialueet ovat tämän työn tarkastelun ulkopuolella.

Työn tarkoituksena on selvittää liittymien eri ominaisuuksien vaikutusta liittymässä tapahtuvien onnettomuuksien lukumäärään. Tämän vuoksi kuljettajien ja ajoneuvojen ominaisuuksiin liittyvät seikat voidaan olettaa eksogeenisiksi, tarkastelun ulkopuolisiksi. Esimerkiksi kuljettajien päihtymystila, ajokokemus tai ajoneuvojen rakenteellinen turvallisuus vaihtelevat samalla tavalla satunnaisesti kaikkialla katuverkossa, eivätkä ne täten vaikuta yksittäisten liittymien keskinäiseen vertailuun. On mahdollista että eri puolilla Suomea liikennekäyttäytymisessä ja ajoneuvokannan turvallisuudessa on alueellisia eroja, mutta ainakin Helsingin sisällä oletus näiden ominaisuuksien tasaisesta jakautumisesta on perusteltu.

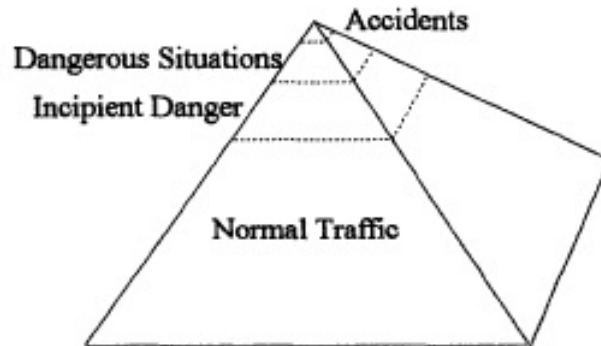
Tutkimuksessa tarkasteltava alue on rajattu Helsingin kantakaupunkiin. Rajauksella pyritään homogenisoimaan tutkimusaineisto siten, että se soveltuu paremmin tilastolliseen mallintamiseen. Tutkimusaineisto ja sen rajaukset esitellään tarkemmin luvussa 4.2.

1.4 Liikenneturvallisuuden määrittäminen

Liikenneturvallisuuden perimmäisenä mittapuuna ovat tyypillisesti liikenneonnettomuuksien seurauksena tapahtuneiden henkilövahinkojen, eli loukkaantumisten ja kuolemantapauksten lukumäärä. Esimerkiksi Suomen tieliikenteen turvallisuusvision tavoitteena on, että kenenkään ei tarvitsisi kuolla tai loukkaantua vakavasti liikenteessä (Liikenne- ja viestintäministeriö 2010). Vastaavasti Euroopan Unionin liikenneturvallisuusohjelmassa on asetettu tavoitteeksi puolittaa liikennekuolemien lukumäärä vuoden 2010 tasosta vuoteen 2020 mennessä (Euroopan komissio 2010).

Handbook of Road Safety Measures -käsikirjassa (Elvik ym. 2009, s. 81) todetaan liikenneturvallisuutta mitattavan tavallisesti tapahtuneiden onnettomuuksien tai henkilövahinkojen lukumääränä. Tämä vastaa esimerkiksi edellä esitettyjä, sekä Liikenne- ja viestintäministeriön että Euroopan unionin asettamia tavoitteita.

Hauer esittää kirjassaan (1997, s. 18-20) kerroksittaisen näkökulman liikenneturvallisuuteen. Tämän viitekehyksen mukaan liikenneturvallisuutta voidaan tarkastella sisäkkäisten osajoukkojen sarjana, tai pyramidina, kuten kuvassa 1.1 on esitetty. Jatkuvasti tapahtuvista liikennetilanteista vain osa johtaa vaaratilanteisiin, joista edelleen vain osa johtaa onnettomuuksiin. Onnettomuuksien joukosta puolestaan henkilövahinko-onnettomuudet ovat vain pieni osajoukko.



Kuva 1.1: Liikenneturvallisuuden tasot (Ezra Hauer 1997)

Hauerin mukaan liikenneturvallisuus ei ole yksiselitteisesti määriteltävissä: esimerkiksi vaaratilanteita ja onnettomuuksia voidaan molempia pitää samalla tavalla huonon liikenneturvallisuuden ilmentyminä. Koska eri tapahtumat ovat vakavuusjärjestyksessä toistensa osajoukkoja, voidaan liikenneturvallisuutta tarkastella yhden osajoukon kautta. Tällä perusteella onnettomuudet ovat hyvä mittari liikenneturvallisuudelle, sillä niiden lukumäärä on selvitetävissä esimerkiksi vaaralliseksi koettujen tilanteiden lukumäärää luotettavammin. Mikäli onnettomuuksien lukumäärää mitataan tiettyinä, tasamittaisina aikaväleinä, käytetään tyypillisesti termiä *onnettomuusfrekvenssi*. (Hauer 1997)

Liikenneturvallisuuden käsikirjassa (Salusjärvi 1992) onnettomuusfrekvenssi esitetään muodostuvan riskistöstä ja riskistä. Riskistö on kirjan määritelmän mukaan onnettomuudelle altistavan toiminnan suuruus, ja riski on yksittäisen onnettomuuden tapahtumistodennäköisyys kyseisen toiminnan yhteydessä, jolloin riskin ja riskistön tulo on onnettomuusfrekvenssi.

Työssä tarkastellusta aiemmasta tutkimuksesta ainoastaan Ezra Hauerin kirjassa (1997, s. 24) on esitetty eksplisiittinen määritelmä liikenneturvallisuudelle. Tässä diplomityössä liikenneturvallisuuden määritelmänä käytetään mukaelmaa tästä määritelmästä:

Liikenneturvallisuus on tietyn ajanjakson aikana tapahtuvien onnettomuuksien odotettu lukumäärä sekä vakavuus

Tässä työssä onnettomuusfrekvenssillä tarkoitetaan yhden vuoden aikana tapahtuvien onnettomuuksien lukumäärää.

Luku 2

Liikenneturvallisuuden mittaaminen

2.1 Mustien pisteiden tunnistaminen

2.1.1 Mustan pisteen määrittely

Liittymien liikenneturvallisuuden parantamiseksi on kyettävä tunnistamaan sellaiset liittymät, joissa liikenneturvallisuus on poikkeuksellisen huono. Tällaisesta kohteesta käytetään usein nimitystä *musta piste*. Mustalle pisteelle (engl. *hot spot* tai *black spot*) ei ole vakiintunutta, yksiselitteistä määritelmää (Geurts & Wets 2003).

Tässä työssä musta piste määritellään *Handbook of Road Safety Measures* -käsikirjaa mukaillen (Elvik ym. 2009, s. 206):

Musta piste on kohde, jossa onnettomuusfrekvenssin odotusarvo on kohteen ominaisuuksista johtuvista syistä muita vastaavia kohteita suurempi

Tämä määritelmä sisältää oletuksen siitä, että kullakin liittymällä on ole-

massa tietty odotusarvo tapahtuvien onnettomuuksien lukumäärälle, eli tietty liittymälle ominainen liikenneturvallisuuden taso. Määritelmään sisältyy myös vaatimus siitä, että onnettomuuksien muita vastaavia kohteita suurempi lukumäärä johtuu kohteen ominaisuuksista.

Liikenneturvallisuuden mittaaminen ja mustien pisteiden tunnistaminen liittyvät läheisesti toisiinsa. Mustien pisteiden tunnistamiseksi on määritettävä sekä eri kohteiden liikenneturvallisuus yhteismitallisella tavalla että menetelmä jolla varsinaiset mustat pisteet voidaan tunnistaa mitatun liikenneturvallisuuden perusteella.

2.1.2 Satunnaisvaihtelun huomioiminen

Onnettomuusmäärille on ominaista voimakas satunnaisvaihtelu, mikä vaikeuttaa liikenneturvallisuuden arviointia etenkin sellaisissa liittymissä joissa tapahtuu tavallisesti vähän onnettomuuksia (Montella 2010, Hauer 1997). Liittymien onnettomuusfrekvenssit ovat tyypillisesti pieniä: tämän työn tutkimusaineiston perusteella Helsingissä yhden liittymän keskimääräinen onnettomuusfrekvenssi on 0,65 onnettomuutta vuodessa. Tällöin yhdelläkin onnettomuudella on merkittävä suhteellinen vaikutus liittymän liikenneturvallisuuteen, mikäli liikenneturvallisuus määritellään yksinkertaisesti onnettomuusfrekvenssinä. Satunnaisvaihtelusta johtuvaa muutosta onnettomuusfrekvenssissä ei kuitenkaan ole syytä pitää muutoksena kohteen varsinaisessa liikenneturvallisuudessa.

Tarkastellaan esimerkkinä liittymää, jossa havaitaan tiettyä vuotena aiempaa enemmän onnettomuuksia. Jos liittymää parannetaan, sen onnettomuusfrekvenssin voidaan olettaa laskevan seuraavana vuonna. Mikäli alkuperäinen muutos onnettomuuksissa on johtunut pelkästä satunnaisvaihtelusta, onnettomuusfrekvenssi todennäköisesti kuitenkin laskisi joka tapauksessa seuraavana vuonna. Tällöin ei ole mahdollista saada selville mikä osuus parannuksesta turvallisuudessa johtuu tehdystä toimenpiteestä, ja mikä osuus pelkästä satunnaisvaihtelusta. Tästä ilmiöstä käytetään englanninkielisessä tutkimus-

kirjallisuudessa yleisesti nimitystä *regression to the mean*. (Hauer 1997)

2.2 Aiempi tutkimus aiheesta

2.2.1 Verrattavien menetelmien esittely

Liikenneturvallisuuden kvantitatiivisesta mittaamisesta ja mustien pisteiden tunnistamismenetelmistä on tehty paljon akateemista tutkimusta, jossa vertaillaan eri menetelmien kykyä identifioida vaaralliset kohteet. Eri tutkimuksissa on vertailtu useita samoja menetelmiä, ja lukemisen helpottamiseksi kirjallisuudessa tärkeimmät tutkimuksissa tarkastellut menetelmät on esitetty alla Montellan (2010) tutkimuksen nimeämiskäytännön mukaisesti.

CF - Crash frequency

Menetelmä on tutkimuksissa verratuista yksinkertaisin: liittymän liikenneturvallisuus on liittymässä havaittu onnettomuusfrekvenssi.

CR - Crash rate

Onnettomuusaste on onnettomuusfrekvenssin ja liikennemäärän suhde. Onnettomuusasteen määrittelyllä pyritään vähentämään altistuksen vaikutusta ja mittaamaan yksittäisen tienkäyttäjän kokemaa riskiä.

EB - Empirical Bayes

Empirical Bayes -luku on painotettu keskiarvo liittymässä havaitusta onnettomuusfrekvenssistä ja liittymän onnettomuusfrekvenssin odotusarvosta. Odotusarvo kuvaa muiden vastaavien liittymien liikenneturvallisuuden tasoa, ja se määritellään tilastollisella mallintamisella. Havaitun onnettomuusfrekvenssin yhdistäminen muiden vastaavien kohteiden onnettomuusfrekvenssiin vähentää satunnaisvaihtelun vaikutusta. Menetelmällä pyritään estimoimaan liittymän todellista liikenneturvallisuutta.

PFI - Potential for improvement

PFI-luku on EB-luvun ja liittymän onnettomuusfrekvenssin odotusarvon erotus. PFI on tulkittavissa liittymän poikkeamaksi odotetusta turvallisuuden tasosta.

2.2.2 Merkittävimmät tutkimukset**Alfonso Montella (2010)**

Tutkimuksessa on verrattu eri menetelmiä liikenneturvallisuuden mittaamiseksi. Menetelmien vertaamiseen on käytetty kolmea eri kriteeriä, jotka mitaavat menetelmien konsistenssia. Konsistenssi on tutkimuksen tapauksessa määritelty tunnistettujen mustien pisteiden joukon yhdenmukaisuudeksi kahden eri vertailuvuoden välillä.

Varsinaiset mustat pisteet on määritetty rajaamalla liittymistä liikenneturvallisuudeltaan vaarallisimmat 1%, 2% tai 10%. Tutkimuksessa tarkastellut menetelmät ovat CF, CR, EB, EB-menetelmä joka huomioi vain henkilövahinko-onnettomuudet, PFI sekä eri onnettomuustyyppien suhteita vertaileva menetelmä. Tutkimusaineistona työssä on ollut erään italialaisen moottoritien onnettomuusaineisto.

Menetelmistä parhaiten suoriutui Empirical Bayes. EB-menetelmän ohella tutkimuksessa verrattiin myös sellaista EB-menetelmää, jonka onnettomuus-

malli on estimoitu vain henkilövahinko-onnettomuuksista. Kaikki onnettomuudet huomioiva EB-menetelmä suoriutui kuitenkin parhaiten mustien pisteiden tunnistamisessa myös siinä tapauksessa kun vertailtavina olivat vain kohteet joissa on tapahtunut henkilövahinko-onnettomuuksia.

Hauer ym. (2004)

Tutkimuksessa on verrattu Empirical Bayes -menetelmän suorituskkyä tilanteissa joissa menetelmän edellyttämä onnettomuusmalli on määritelty eri kriteerein. Tutkittuja kriteerejä ovat esimerkiksi kaikki onnettomuudet, vakavuudella painotetut onnettomuudet sekä keskimääräisen tason ylittävät onnettomuudet. Tutkimusaineistona käytetään liittymiä, joissa on tehty parannustoimenpiteitä, ja joista voidaan tehdä liikenneturvallisuuden ennen-jälkeen-analyysi. Tutkimuksessa havaitaan, että Empirical Bayes suoriutuu parhaiten siinä tapauksessa kun käytetty onnettomuusmalli on estimoitu kaikista onnettomuuksista ilman onnettomuuksien vakavuuteen liittyvää painotusta.

Persaud ym. (2010)

Tutkimuksessa on verrattu Empirical Bayes -menetelmän sekä täysin bayesilaisen menetelmän soveltuvuutta liikenneturvallisuutta parantavien toimenpiteiden ennen-jälkeen-vertailuun. Menetelmien välillä ei havaita eroa, ja tutkijat pitävät yksinkertaisempaa Empirical Bayes -menetelmää riittävänä liikenneturvallisuuden arvioinnissa.

Cheng & Washington (2005)

Tutkimuksessa on verrattavana kolme eri tunnistamismenetelmää: CF, EB ja luottamusväleihin perustuva menetelmä. Tutkimusaineistona on simuloitu onnettomuusdata, jota tutkimuksen tekijät pitävät ylivoimaisena empiiriseen aineistoon verrattuna. Empirical Bayes osoittautui tutkimuksessa selvästi muita verrattuja menetelmiä paremmaksi.

Geurts & Wets (2003)

Kirjallisuustutkimuksessa tarkastellaan mustien pisteiden tunnistamiskäytäntöjä eri maissa sekä laaditaan laaja katsaus aiheesta tehtyyn aiempaan tutkimukseen. Tämä on ainut tässä diplomityössä käytetyistä tutkimuksissa, jossa on mukana nykyisin Helsingissä ja muualla Suomessa käytetyn IND5-menetelmän tapainen henkilövahinkoja painottava mittari. Belgiassa on käytössä menetelmä, jossa huomioidaan eri painokertoimilla lievät loukkaantumiset, vakavat loukkaantumiset sekä kuolemat. Tutkimuksen lopputuloksena on kuitenkin suositus Empirical Bayes -menetelmän käytöstä.

Kusumawati & Wong (2010)

Tutkimuksessa Singaporen mitataan liikennevalo-ohjauksisten liittymien liikenneturvallisuutta ja verrataan kahden eri menetelmän kykyä tunnistaa turvallisuudeltaan heikoimmat kohteet. Musta piste määritellään sellaiseksi liittymäksi, jossa EB-luku on tilastollisesti merkitsevällä tavalla suurempi kuin onnettomuusmallin ennustama onnettomuusfrekvenssi. Verrattavat menetelmät ovat edellä esitelty PFI sekä toinen vastaava menetelmä, joka lasketaan liittymän EB-luvun ja onnettomuusmallin odotusarvon suhteena. Tutkimuksen johtopäätöksenä suositellaan PFI:n käyttämistä.

Kulmala (1995)

Risto Kulmalan väitöskirjassa on tutkittu Suomen maanteiden kolmi- ja neliahaaraliittymien liikenneturvallisuutta Empirical Bayes -menetelmällä. Tutkimuksessa on laadittu erilliset onnettomuusmallit kolmi- ja neliaharaisille liittymille. Mallien selittävinä muuttujina on käytetty muun muassa liittymän liikennemäärää, valaistusta, päällysteen leveyttä, mäkisyyttä sekä liikennevalo-ohjausta. Kulmalan havaintojen mukaan onnettomuusmalleilla voidaan selittää onnettomuuksien lukumäärää liittymissä erittäin hyvin, mutta tutkimuksessa kuvaillaan myös selittävien muuttujien korrelaatiosta johtuvia ongelmia.

2.2.3 Johtopäätökset kirjallisuudesta

Empirical Bayes on mukana lähes kaikissa tutkimuksissa, ja sitä pidetään yleisesti parhaana menetelmänä liikenneturvallisuuden mittaamiseen. Empirical Bayes yhdistää liittymässä havaitun onnettomuusfrekvenssin ja muilta vastaavilta kohteilta odotetun onnettomuusfrekvenssin, mikä vähentää onnettomuuksien voimakkaan satunnaisvaihtelun vaikutusta.

Varsinaiseen mustien pisteiden määrittelyyn tutkimuksissa on kaksi eri lähestymistapaa. Osassa tarkastelluista tutkimuksista mustat pisteet määritetään järjestämällä kohteet liikenneturvallisuuden mukaiseen laskevaan järjestykseen, jolloin vaarallisimmat kohteet ovat järjestyksessä ensimmäisinä. Osassa puolestaan määritetään tietty raja-arvo, joka erottaa vaaralliset kohteet muista.

Sekä Hauerin (2004) että Montellan (2010) tutkimuksissa on havaittu, että Empirical Bayes -menetelmä tunnistaa vaaralliset kohteet parhaiten silloin, kun käytettävä onnettomuusmalli on estimoitu aineistosta joka kattaa kaikki onnettomuudet. Kummassakin tutkimuksessa tällainen malli suoriutui vertailussa paremmin kuin malli, jonka estimoinnissa on painotettu henkilövahinko-onnettomuuksia.

2.2.4 Empirical Bayes

Empirical Bayes -menetelmä yhdistää havaitun onnettomuusfrekvenssin y sekä vastaavilta kohteilta odotetun onnettomuusfrekvenssin $E(Y)$. Odotusarvo $E(Y)$ määritetään onnettomuusmallin avulla. Onnettomuusmalli on tilastollinen malli, jolla pyritään ennustamaan liittymän onnettomuusfrekvenssiä tiettyjen tunnettujen parametrien funktiona. Malli estimoidaan empiirisen havaintoaineiston pohjalta, ja sen laatimista käsitellään tarkemmin luvussa 3.2.

EB-luku on painotettu keskiarvo havaitusta onnettomuusfrekvenssistä ja mallin ennustamasta onnettomuusfrekvenssistä (Hauer 1997):

$$EB = wE(Y) + (1 - w)y \quad (2.1)$$

jossa painokerroin w on

$$w = \frac{1}{1 + \frac{VAR(Y)}{E(Y)}} \quad (2.2)$$

Painokerroin voidaan myös esittää muodossa

$$w = \frac{1}{1 + \frac{E(Y)}{k}} \quad (2.3)$$

jossa k on onnettomuuksien odotusarvoa ennustavan mallin hajontaparametrin käänteisluku (Miaou & Lord 2005).

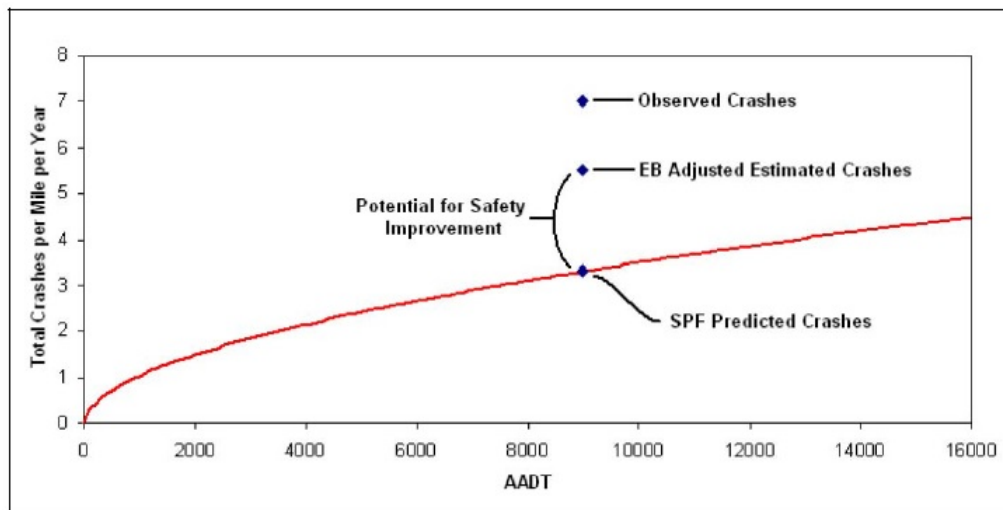
Painokertoimen w esitysmuodosta 2.2 voidaan havaita, että mallin tarkkuus vaikuttaa siihen painotetaanko estimaatissa enemmän havaittua vai ennustettua frekvenssiä. Suuri varianssi johtaa lukuarvoltaan pieneen painokertoimeen ja täten havaitun frekvenssin suurempaan painoarvoon. Tämä havaintojen ja odotusarvon painottaminen vähentää satunnaisvaihtelua ja huomioi onnettomuusfrekvenssin regression keskiarvoa kohti (Geurts & Wets 2003).

Erikoistapauksena menetelmässä on puhdas Poisson-malli, jonka keskiarvo ja varianssi ovat yhtä suuret. Tässä tapauksessa painokerroin w saa arvon 0,5, jolloin sekä havaittu frekvenssi että mallin antama odotusarvo ovat samanarvoisia EB-estimaatissa.

Estimaatin tulkinta

EB on estimaatti liittymän turvallisuudelle, ja sen mittayksikkönä on onnettomuusfrekvenssi, eli tyypillisesti onnettomuuksien lukumäärä vuodessa. EB on tulkittavissa liittymän turvallisuuden tasoksi, eli siksi onnettomuusfrekvenssiksi joka on liittymälle ominainen. Kuvassa 2.1 on esitys EB-

estimaatin, onnettomuusmallin ja havaitun onnettomuusfrekvenssin suhteesta (Virginia DOT 2010). Mikäli EB-estimaatti tulkitaan liittymän turvallisuuden tasoksi ja onnettomuusmallin tuottama arvo sellaiseksi turvallisuuden tasoksi, joka liittymältä voidaan olettaa, on tällöin näiden erotus se parannus turvallisuudessa joka liittymässä voidaan eri toimenpitein saavuttaa. Tämä parannuspotentiaali on edellä mainittu PFI, Potential For Improvement.



Kuva 2.1: EB-estimaatin graafinen tulkinta (Virginia DOT 2010)

Menetelmällä on myös sen nimen mukaisesti bayesilainen tulkinta. Empirical Bayes yhdistää *a priori* -arvion liittymän liikenneturvallisuudesta *a posteriori* -arvioon. *A priori* on tässä tapauksessa onnettomuusmallin tuottama ennuste liittymän onnettomuusfrekvenssistä, ja *a posteriori* on liittymässä havaittu onnettomuusfrekvenssi (Persaud ym. 2010).

2.3 Onnettomuuksien satunnaisluonne

2.3.1 Onnettomuuksien todennäköisyysjakauma

Liikenneonnettomuuksien syntyä voidaan tarkastella satunnaisprosessina, jolloin yksittäisen onnettomuuden tapahtuminen on luonteeltaan toistokoe. Ajoneuvon ajaessa liittymään sillä on tietty todennäköisyys p joutua onnettomuuteen. Vastaavasti todennäköisyys olla joutumatta onnettomuuteen on $q = 1 - p$. Lord käsittelee tutkimuksessaan (2004) onnettomuuksien satunnaisluonnetta, ja osoittaa, että tämänkaltaista tilannetta voidaan approksimoida Poisson-jakaumalla kun toistojen lukumäärä on suuri. Poisson-jakauman pistetodennäköisyysfunktio, eli todennäköisyys sille että onnettomuuksien lukumäärä *tietyssä liittymässä* on y havaintojakson aikana:

$$Pr(Y = y) = \frac{\lambda^y}{y!} e^{-\lambda} \quad (2.4)$$

Termi λ on Poisson-jakauman intensiteetti, eli tietyssä ajanjaksossa esiintyvien tapahtumien lukumäärän odotusarvo.

Poisson-jakauman käyttöä voidaan perustella myös tarkastelemalla onnettomuuksien syntyä yhdessä liittymässä Poisson-prosessina. Mikrotason prosessit, tässä tapauksessa yksittäisten ajoneuvojen saapumiset liittymään, generoivat vain harvoin onnettomuuden. Näillä mikrotason prosesseilla ei välttämättä ole Poisson-luonnetta, vaan onnettomuuden todennäköisyys voi vaihdella jokaisen kuljettaja-ajoneuvo-olosuhde-kombinaation mukaan.

Palm-Khintchine-teoreeman mukaan nämä keskenään heterogeeniset prosessit kuitenkin käyttäytyvät makrotasolla Poisson-prosessin omaisesti mikäli toistojen määrä on suuri. (Tijms 2003, s. 2)

Oletus Poisson-jakaumasta voidaan laajentaa koskemaan kaikkia liittymiä siten, että yksittäisen liittymän i onnettomuuksien lukumäärä on Poisson-jakautunut odotusarvolla $\lambda = \eta_i \theta$. Keskiarvojen vaihtelua liittymien välillä kuvaava parametri θ on gamma-jakautunut odotusarvolla $E(\theta) = 1$ ja va-

rianssilla $Var(\theta) = 1/\phi$, jossa ϕ on jakauman hajonta- tai dispersioparametri. Tällöin onnettomuuksien lukumäärä liittymässä i , Y_i , noudattaa negatiivista binomijakaumaa (Lord 2006, Hauer 2001, Kulmala & Roine 1990). Jakauman pistetodennäköisyysfunktio on:

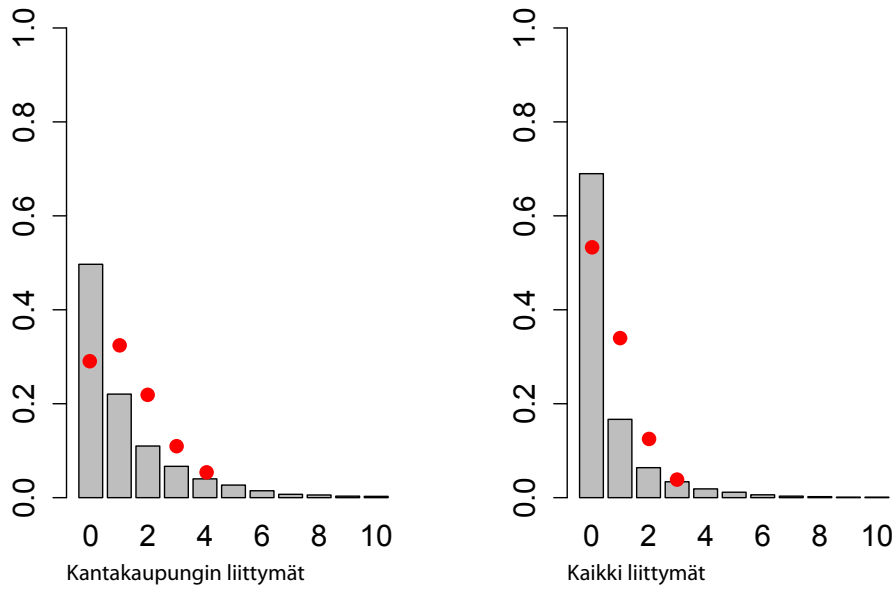
$$Pr(Y_i = y_i) = \frac{\Gamma(y_i + \phi)}{\Gamma(\phi)y_i!} \left(\frac{\phi}{\eta_i + \phi} \right)^\phi \left(\frac{\eta_i}{\eta_i + \phi} \right)^{y_i} \quad (2.5)$$

2.3.2 Ylihajonta

Poisson-jakauman varianssi ja keskiarvo ovat yhtä suuret, mikä ei kuitenkaan usein pidä paikkaansa empiirisen aineiston kanssa. Onnettomuusaineistolle on tunnuksenomaista ylihajonta, eli pelkän Poisson-mallin ennustamaa suurempi varianssi. Erityisesti ongelmana on usein nollien suuri lukumäärä. (Esim. Lord 2006a, Hauer 2001).

Kuvassa 2.2 on esitetty histogrammit kaikkien Helsingin liittymien sekä kantakaupungin liittymien onnettomuusfrekvensseistä vuosina 2005–2011. Havaintoaineiston päälle on sovitettu Poisson-jakauman pistetodennäköisyysfunktion arvot siten, että Poisson-jakauman intensiteetti (keskiarvo) λ on sama kuin onnettomuuksien lukumäärän keskiarvo. Histogrammien pylväät kuvaavat eri onnettomuusfrekvenssien (0, 1, 2...) suhteellista osuutta havaintoaineistossa. Pylväiden päällä olevat pisteet kuvaavat sitä suhteellista osuutta, joka kullakin onnettomuuskategorialla tulisi olla mikäli havaitut onnettomuudet olisivat täysin Poisson-jakautuneita.

Ylihajonta on selvästi havaittavissa kuvasta: sellaisia liittymiä joissa onnettomuuksia ei ole tapahtunut on suurempi määrä kuin Poisson-jakauman tapauksessa.



Kuva 2.2: Ylihajonta tutkimusaineistossa

Yksi mahdollisuus ylihajonnan huomiointiin on negatiivisen binomijakauman käyttö. Edellä esitelty malli on kaksiparametrinen, ja keskiarvon lisäksi estimoitava hajontaparametri mahdollistaa keskiarvoa suuremman varianssin.

Myös Poisson-jakautuneelle muuttujalle Y , jolla $E(Y) = \mu$, on mahdollista merkitä varianssi siten, että $var(Y) = \phi\mu$. Hajontaparametri ϕ saa arvon 1 aidon Poisson-jakauman tapauksessa. Mikäli Poisson-jakauman varianssi merkitään hajontaparametrin avulla, mallin parametrejä ei estimoida suurimman uskottavuuden menetelmällä vaan kvasiuskottavuuden menetelmällä. Mallin parametrien kannalta tällä ei ole suurta merkitystä, mutta mallille ei voida määritellä tiettyjä uskottavuusfunktioon perustuvia arvoja, kuten mallien keskinäisessä vertailussa käytettävää Akaiken informaatiokriteeriä. (Hilbe 2011 s. 164, Venables & Ripley 2002 s. 186-210)

Luku 3

Tutkimusmenetelmät

3.1 Mustien pisteiden tunnistaminen tutkimuksessa

Tässä työssä liittymän liikenneturvallisuuden selvittämiseen ja mustien pisteiden tunnistamiseen käytetään Empirical Bayes-menetelmää. Luvussa 2.2.1 esitellyn aiemman tutkimuksen perusteella Empirical Bayes -menetelmä on eri menetelmistä käytetyin ja luotettavin. Empirical Bayes -menetelmän soveltamista varten työssä laaditaan onnettomuusmalli, jolla ennustetaan liittymän onnettomuusfrekvenssin odotusarvoa. Tämä mallin tuottama tulos sekä liittymässä havaittu onnettomuusfrekvenssi yhdistetään estimaatiksi liittymän liikenneturvallisuudesta. EB-luku lasketaan kullekin liittymälle kaavan 5.2.2 mukaisesti.

Varsinaiset mustat pisteet määritetään järjestämällä liittymät jokaisena tarkasteluvuotena Empirical Bayes -luvun mukaan laskevaan järjestykseen. Työssä ei käytetä mitään tiettyä raja-arvoa erottamaan poikkeuksellisen vaaralliset liittymät vähemmän vaarallisista, vaan liittymiä arvioidaan ensisijaisesti niiden keskinäisen järjestyksen perusteella.

Empirical Bayes -luvun lisäksi liittymille määritetään Potential For Improve-

ment, eli PFI-luku. PFI on EB-luvun ja mallin ennustaman onnettomuusfrekvenssin odotusarvon erotus, ja kuvaa liittymän liikenneturvallisuuden poikkeamaa muista vastaavista liittymistä. Tämä luku auttaa yksittäisten liittymien arvioinnissa ja tarjoaa lisäinformaatiota liittymästä suhteessa muihin samanlaisiin liittymiin. PFI kuvaa periaatetasolla myös potentiaalista parannusta liikenneturvallisuudessa, joka liittymässä voidaan eri toimenpitein saavuttaa.

3.2 Onnettomuuksien mallintaminen

3.2.1 Yleistetty lineaarinen malli

Empirical Bayes -menetelmän käyttäminen edellyttää onnettomuusmallin laatimista. Onnettomuuksien mallintamisen tavoitteena on löytää sellainen tilastollinen malli, jolla voidaan selittää onnettomuuksien lukumäärää tiettyjen tunnettujen parametrien funktiona.

Onnettomuudet ja yleisesti lukumääräiset vastemuuttujat eivät sovellu mallinnettavaksi tavallisella lineaarisella regressiolla, sillä lineaarisessa regressiossa jäännöstermien oletetaan olevan normaalijakautuneita. Mikäli vastemuuttuja voi saada vain positiivisia arvoja, jäännöstermi ei voi olla normaalijakautunut pienillä vastemuuttujan odotusarvoilla. Tällaisessa tapauksessa yleistetty lineaarinen malli, jonka erikoistapaus lineaarinen regressiomalli on, toimii luotettavammin. (Heikkinen 2005, Kusumawati & Wong 2010)

Yleistetyn lineaarisen mallin tapauksessa riippuvuus vastemuuttujan ja selittävien muuttujien välillä ei ole välttämättä lineaarinen, vaan muuttujat liittyvät toisiinsa linkkifunktion avulla. Olkoon satunnaismuuttujan Y_i odotusarvo $E(Y_i) = \mu_i$. Tällöin yleistetty lineaarinen malli on muotoa (Kulmala & Roine 1990, s. 28):

$$g(\mu) = \alpha + \beta_1 x_{i1} + \dots + \beta_p x_{ip} \quad (3.1)$$

jossa α on mallin vakiotermi, x on selittävä muuttuja ja β selittävää muuttujaa vastaava regressiokerroin.

Funktio $g(\mu)$ on linkkifunktio, joka kuvaa vastemuuttujan muuttujan arvot selitettävien muuttujien arvojen lineaarikombinaatioksi (Heikkinen 2005, s 14). Vastaavasti linkkifunktion käänteisfunktio g^{-1} kuvaa selittävien muuttujien lineaarikombinaation vastemuuttujan arvoiksi. Linkkifunktio määräytyy useimmissa tapauksissa vastemuuttujan todennäköisyysjakaumasta. Eri jakaumille on määritelty tietyt oletusarvoisesti käytettävät linkkifunktiot, joista käytetään nimitystä *kanoninen linkkifunktio*. (Mellin 2011)

Mallintaminen ja kaikki aineiston käsittely on toteutettu R:llä, avoimen lähdekoodin tilastollisella ohjelmointikielellä. Yleistetty lineaarinen malli on toteutettu R-kieleen MASS-nimisellä laajennuksella. Laajennus on Modern Applied Statistics with S -kirjan tekijöiden (Venables & Ripley 2002) laatima, ja se sisältää yleistettyjen lineaaristen mallien estimoimiseksi vaadittavat funktiot.

3.2.2 Mallin jakaumaoletus

Työssä käytetään onnettomuuksien mallintamiseen negatiivista binomijakaumaa. Se on helposti käytettävissä nykyaikaisissa tilastotieteen ohjelmistoissa, ja tarjoaa kirjallisuuden mukaan paremman sovituksen empiiriseen dataan kuin pelkkä Poisson-malli. Erään kirjallisuustutkimuksen (Geurts & Wets 2003) mukaan negatiivinen binomimalli on käytetyin ja yleisesti parhaana pidetty jakaumaoletus onnettomuuksille. Luvussa 2.3.1 on perusteltu tarkemmin kuinka onnettomuuksien tapahtumista voidaan kuvata negatiivisella binomijakaumalla.

3.2.3 Mallin rakenteen valinta

Mallin rakenteen osalta määritellään käytetty linkkifunktio sekä mallissa käytettävät selittävät muuttujat. Kuten luvussa 3.2.2 on todettu, tässä työssä onnettomuuksien lukumäärän oletetaan noudattavan negatiivista binomija-

kaumaa. Negatiivisen binomijakauman tapauksessa käytettävä linkkifunktio on tyypillisesti logaritminen (Ver Hoef & Boveng 2007). Tällöin vastemuuttuja voidaan esittää kaavan 3.1 mukaisesti muodossa:

$$\mu = e^{\alpha + \beta_1 x_{i1} + \dots + \beta_p x_{ip}} \quad (3.2)$$

Käytetyssä kirjallisuudessa on tunnistettavissa kaksi erilaista lähestymistapaa mallin rakenteeseen. Suuressa osassa työssä käytetystä akateemisesta tutkimuksesta malli on rakennettu siten, että kaikki eri onnettomuuksiin vaikuttavat tekijät ovat omina selittävinä muuttujinaan. Esimerkkejä tätä menetelytapaa edustavista tutkimuksista ovat Montella (2010), Mountain & Fawaz (1996) sekä Kulmala & Roine (1990). Tämänkaltaisen malli mahdollistaa saman mallin käyttämisen kaikille eri liittymille. Heikkoutena menetelmässä on vaatimus tarkasta mallintamisaineistosta, joka on edellytys muuttujien käyttämiselle mallintamisessa. Lisäksi mallin ylihajonnan huomioiva dispersioparametri on tällöin sama kaikille liittymille.

Toinen mahdollisuus on sijoittaa liittymät tiettyjen muuttujien perusteella eri kategorioihin ja estimoida jokaiselle kategorialle oma mallinsa. Highway Safety Manual noudattaa tätä tapaa, ja menetelmää käytetäänkin Yhdysvalloissa. Esimerkiksi Virginian osavaltion liikenneviranomaisen käyttämissä onnettomuusmalleissa liittymät on jaettu eri kategorioihin muun muassa haarojen lukumäärän, liikennevalojen sekä STOP-merkkien mukaan (Virginia Department of Transportation 2010). Eräitä kategorioita ovat ”maaseudun kolmihaaraliittymä sivusuunnan STOP-merkeillä” sekä ”kaupungin nelihhaaraliittymä valo-ohjauksella”. Ainoat selittävät muuttujat malleissa ovat vakiotermin lisäksi pää- ja sivusuunnan liikennemäärät.

Tämänkaltaisen malli on helpompi käyttää parametrien yksinkertaisuuden vuoksi, ja mahdollistaa dispersioparametrin estimoinnin yksilöllisesti eri liittymätyypeille. Heikkoutena tässä lähestymistavassa on mahdollinen mallin huonompi sopivuus aineistoon, mikäli mallin sovitusta voitaisiin parantaa selittäviä muuttujia lisäämällä. Lisäksi mikäli kategorioita on paljon, saattaa havaintojen lukumäärä osassa liittymätyypeistä olla liian pieni mallintamista

varten. Esimerkiksi Virginian tapauksessa havaintoaineistoltaan laajimmassa kategoriassa oli 2114 havaintoa ja pienimmässä vain 33.

Miaou ja Lord (2003) esittävät tutkimuksessaan mahdollisuuden määrittää hajontaparametri liittymän ominaisuuksien funktiona. Tällöin käytettyjen mallien lukumäärä ei rajoita hajontaparametrin tarkkuutta, eikä tarvetta erillisille mallikategorioille ole. Lähestymistavasta ei kuitenkaan ole saatavissa riittävää tietoa eikä sille ole riittävästi tukea käytetyissä tilastotieteen ohjelmistoissa, jotta sitä voitaisiin hyödyntää tässä tutkimuksessa.

Työssä käytettävä mallin rakenne

Tässä työssä mallin rakenteena käytetään yhdistelmää yllä esitetyistä. Liittymät jaetaan kahteen kategoriaan jäljempänä perusteltavan, dikotomisen muuttujan suhteen. Mallissa käytetään kuitenkin pelkkien liikennemäärien lisäksi muitakin selittäviä muuttujia. Tämä tarjoaa tasapainon mallintamisen kattavuuden sekä käytön helppouden välillä. Esimerkiksi Singaporessa tehdyssä tutkimuksessa (Kusumawati & Wong 2010) käytetty onnettomuusmalli on esitetyn kaltainen: kolmi- ja nelihaaraisille liittymille on omat mallinsa, ja muuttujina ovat liikennemäärien lisäksi liittymän katuluokka sekä liikenneympäristön kaupunkimaisuus.

3.2.4 Vastemuuttujan valinta

Onnettomuusmallilla pyritään ennustamaan liittymän liikenneturvallisuutta, jota mitataan onnettomuusfrekvenssillä. Onnettomuusfrekvenssin lukuarvo voidaan kuitenkin määrittää eri tavoin käytetystä aineistosta riippuen.

Käytettävä onnettomuuksien lukumäärä

On mahdotonta saada täysin tarkkaa tietoa siitä kuinka monta liikenneonnettomuutta Suomessa todellisuudessa tapahtuu. Liikenneturvan ja Tilasto-

keskuksen julkaiseman virallisen onnettomuustilaston mukaan vuonna 2012 tapahtui yhteensä 33 065 poliisin tietoon tullutta onnettomuutta, joista 5 490 johti loukkaantumiseen ja 235 kuolemaan. Kuitenkin samana vuonna Liikennevakuutuskeskus on tilastoinut kaikkiaan 113 000 sellaista onnettomuutta, joiden vahinkoja on korvattu liikennevakuutuksesta (Tilastokeskus 2015).

Edellä esitetyistä luvuista on pääteltävissä, että kaikki onnettomuudet eivät tule poliisin tietoon. Sellaiset onnettomuudet, joista ei ole ilmoitettu poliisille tai joihin ei ole haettu korvausta liikennevakuutuksesta eivät ole mukana missään tilastoluvussa, eikä näiden onnettomuuksien lukumäärän selvittäminen ole mahdollista. Vakavat henkilövahinko-onnettomuudet kuitenkin ilmoitetaan poliiseille todennäköisemmin kuin pelkät omaisuusvahinko-onnettomuudet, jolloin ne päätyvät myös suuremmalla todennäköisyydellä onnettomuustilastoon. Liikenneturvan (2012) mukaan kaikki kuolemaan johtaneet onnettomuudet ovat mukana virallisessa tieliikenneonnettomuustilastossa, mutta loukkaantumiseen johtaneista onnettomuuksista arviolta vain 20 prosenttia.

Tilastokeskuksen (2015) käyttämän määritelmän mukaan loukkaantuneiksi luokitellaan sellaiset henkilöt joiden vammat vaativat hoitoa. Sama loukkaantumisen kategoria käsittää siis kaikki vammat lievistä luunmurtumasta elinikäiseen liikuntakyvyttömyyteen. Tämä on yksi syy siihen miksi onnettomuuksien vakavuutta on vaikea arvioida helposti saatavien tilastolukujen perusteella.

Jotta käytetty aineisto on mahdollisimman yhdenmukaista ja samoin perustein kerättyä, tässä työssä onnettomuuksien lukumääränä käytetään kaikkien poliisin tietoon tulleiden onnettomuuksien lukumäärää. Onnettomuusaineisto esitellään tarkemmin luvussa 4.1.1.

Onnettomuuksien vakavuuden huomiointi

Liikenneturvallisuutta arvioitaessa henkilövahingoilla on suurempi merkitys kuin pelkillä omaisuusvahinko-onnettomuuksilla. Onnettomuuksien vakavuuden huomioimiseksi mallintamisessa ei kuitenkaan ole aiemmassa tutkimuksessa tunnistettavissa mitään vakiintunutta, yleisesti käytettyä menetelmää. Yksi mahdollisuus on huomioida henkilövahingot jollain tavoin mallin rakenteessa, esimerkiksi vastemuuttujan valinnalla.

Kulmalan ja Roineen (1990) tutkimuksessa on käytetty mallintamisessa vastemuuttujana henkilövahinko-onnettomuuksia, ja valintaa on perusteltu henkilövahinko-onnettomuuksien tilastoinnin paremmalla luotettavuudella. Yhdysvalloissa Virginian osavaltion liikenneviranomaisen (Virginia DOT 2010) käyttämässä ohjeistuksessa on laadittu erilliset mallit kaikille onnettomuuksille sekä henkilövahinko-onnettomuuksille.

Luvussa 2.2.1 esitellyissä tutkimuksissa on kuitenkin havaittu kaikkien onnettomuuksien olevan vakavuudella painotettua onnettomuuslukua parempi vastemuuttuja onnettomuusmallissa. Tämän perusteella tässä työssä mallintamisen vastemuuttujana käytetään kaikkien onnettomuuksien lukumäärää. Tämä ei tarkoita että onnettomuuksien vakavuudella ei olisi merkitystä. Tutkimusten mukaan kaikkien onnettomuuksien käyttäminen ennustamisessa auttaa tunnistamaan myös henkilövahinko-onnettomuuksissa mitattuna vaarallisimmat kohteet muita menetelmiä luotettavammin.

3.3 Mallien arviointi

3.3.1 Multikollineaarisuus

Lineaarisissa malleissa selittävät muuttujat oletetaan keskenään riippumattomiksi, mutta empiirisessä aineistossa oletus ei tyypillisesti ole realistinen. Usean selittäjän lineaarisessa mallissa selittävien muuttujien välisestä korre-

laatiosta käytetään nimitystä *multikollineaarisuus*.

Multikollineaarisuus ei heikennä mallin selityssastetta tai mallin ennustuskkyä, mutta se vaikeuttaa yksittäisten muuttujien vaikutuksen arviointia. Mikäli esimerkiksi liittymän pää- ja sivusuuntien liikennemäärien välillä on voimakas positiivinen korrelaatio, ei pelkän pääsuunnan liikennemäärän vaikutuksen suuruutta voida arvioida mallin pohjalta. (Gujarati & Porter 2009, s. 253)

Tässä työssä mallissa käytettävien selittävien muuttujien välisestä korrelaatiosta johtuvan multikollineaarisuuden määrään ei voida vaikuttaa, mutta sen olemassaolo huomioidaan laadittavista malleista tehtävissä tulkinnoissa.

3.3.2 Mallien tilastollinen merkitsevyys

Lineaarisen regressiomallin hyvyttä mitataan selityssasteella, josta käytetään merkintää R^2 . Selityssaste vaihtelee välillä 0...1, ja se on tulkittavissa mallin selittämänä osuutena aineiston kokonaisvaihtelusta. Esimerkiksi selityssasteen 0,9 regressiomalli selittää 90 % aineiston vaihtelusta. (Gujarati & Porter 2009)

Yleistetyille lineaarisille malleille ei kuitenkaan ole vastaavaa, mallin kontekstista riippumatonta hyvyyden mittaria. Yleistettyjä lineaarisia malleja voidaan vertailla useilla eri kriteereillä, mutta niille kaikille on yhtenäistä se, että ne soveltuvat vain samasta aineistosta estimoitujen mallien keskinäiseen vertailuun. (Hilbe 2011 s. 65)

Esimerkkejä onnettomuusmalleille käytetyistä hyvyyden mittareista ovat mean deviance ratio (MDR) (Mountain & Fawaz 1996) sekä scaled deviance (Kulmala & Roine 1990). Molemmat menetelmät vertaavat tarkasteltavan mallin hajontaa samasta aineistosta estimoidun, yksinkertaisemman mallin hajontaan.

Tässä työssä mallien keskinäisen hyvyyden vertailussa käytettävä menetelmä on Akaiken informaatiokriteeri, (AIC). Samasta aineistosta estimoituja mal-

leja verrattaessa pienempi AIC:n arvo tarkoittaa mallin sopivan aineistoon paremmin. AIC:n lukuarvolla ei ole merkitystä, mutta kahden mallin informaatiokriteereiden erotus on merkittävä. Erotus jonka suuruus on 2 tarkoittaa mallien eroavan toisistaan tilastollisesti merkitsevällä tavalla, ja erotus jonka suuruus on 10 on merkki siitä että toinen malli sopii aineistoon varmasti paremmin (Burnham & Anderson 2002). AIC voidaan laskea kaavalla:

$$AIC = 2k - 2\ln(L) \quad (3.3)$$

Yhtälössä k on mallin parametrien lukumäärä ja L on mallin uskottavuusfunktion suurin arvo.

3.4 Tulosten arviointi

Työssä arvioidaan myös onko Empirical Bayes -menetelmä nykyistä IND5-lukua parempi keino liittymien liikenneturvallisuuden mittaamiseen. Menetelmien eroa arvioidaan testillä, joka on mukautettu Montellan tutkimukseen (2010) käyttämistä menetelmistä.

Käytettävällä testillä pyritään mittaamaan eri menetelmien kykyä tunnistaa vaaralliset kohteet johdonmukaisesti eri vuosien välillä. Mikäli liikenneturvallisuus on liittymän ominaisuus, eikä siinä tapahtu muutosta eri vuosien välillä, ideaalitulanteessa mustien pisteiden joukon tulisi olla vuodesta toiseen sama.

Testissä verrataan perusvuoden kahdenkymmenen vaarallisimman liittymän joukkoa vertailuvuosien kahdenkymmenen vaarallisimman liittymän joukkoon. Testisuureen arvo saadaan laskemalla kuinka moni vuoden 2011 kahdestakymmenestä vaarallisimmasta liittymästä esiintyy vertailuvuosien kahdenkymmenen vaarallisimman kohteen joukossa. Mitä suurempi testisuureen arvo on, sitä konsistentimpi on vaarallisten liittymien joukko eri vuosien välillä. Testisuure lasketaan IND5-, EB- sekä PFI-menetelmille. Perusvuote-

na testissä on 2011, ja vertailuvuosina 2010 sekä 2009.

Luku 4

Tutkimusaineisto

4.1 Käytetyt aineistot

4.1.1 Onnettomuusaineisto

Työssä käytetty onnettomuusaineisto koostuu poliisin tietoon tulleista ja Helsingin kaupunkisuunnitteluviraston tilastoimista onnettomuuksista. Aineisto kattaa onnettomuudet vuosilta 2003-2011, ja yhteensä havaintoja on n. 32 000.

Helsingin poliisilaitos toimittaa Helsingin kaupungille määrämuotoisille lomakkeille tallennetut tiedot tapahtuneista onnettomuuksista. Lomakkeet sisältävät tiedot osallisista, kuvauksen onnettomuudesta ja piirroksen onnettomuustapahtumasta. Tiedot tallennetaan Helsingin kaupungin onnettomuusrekisteriin, johon tallennetaan myös tiedot osallisten kulkusuunnista sekä tieto onnettomuuden sijainnista. Onnettomuudet kirjataan vain mikäli niiden tapahtumapaikka on tarkasti tiedossa. Rekisteriä täydennetään tiedoilla onnettomuuspaikasta Helsingin omien katuverkkoa kuvaavien rekisterien avulla sekä Liikenneviraston kautta saatavilla tiedoilla onnettomuuksien tapahtumisolosuhteista.

Onnettomuusrekisterissä on tiedot muun muassa onnettomuuden tapahtumajasta sekä -paikasta, säätilasta sekä liikennevalojen toiminnasta.

4.1.2 Liittymäpiste- ja liikennemääräaineisto

Toinen pääasiallinen aineistolähde on liittymäpisterekisteri, joka sisältää Helsingin alueella sijaitsevien liittymien tiedot haarakohtaisesti. Jokaisella liittymällä on aineistossa tyypillisesti kolme tai neljä haaraa siten, että haarat 1 ja 3 sijaitsevat vastakkaisilla puolilla liittymää. Kolmihaaraisella liittymällä vastinpariton haara on haara numero 2, ja nelihaaraisen liittymän tapauksessa haarat 2 ja 4 ovat vastakkaisilla puolilla. Liittymäpisterekisteriin on taltioitu muun muassa liittymän haaran nopeusrajoitus sekä kadun nimi.

Liittymien identifiointiin käytetään ruutupiste-koodia: ruutu on se karttaruutu jossa liittymä sijaitsee ja piste on yhden ruudun sisällä sijaitsevien liittymien juokseva numerointi. Esimerkkinä G1P302, jossa G1P3 on karttaruutu ja 02 liittymäpisteen numero. Liittymien nimeäminen on kuitenkin muuttunut tämän työn laatimisen aikana, ja tulevaisuudessa liittymiin viitataan juoksevalla numeroinnilla. Työn tutkimusaineisto on tarvittaessa muutettavissa uuden nimeämiskäytännön mukaiseksi.

Liittymäpisterekisteriin on kirjattu sellaiset liittymät, joissa on joskus tapahtunut onnettomuus. Tämä aiheuttaa lievää vääristymää mallintamiseen, sillä sellaiset liittymät joissa ei koskaan ole tapahtunut onnettomuuksia eivät ole mukana aineistossa. Liittymäaineisto kuitenkin on kerätty tämän työn tarkastelujaksoa pidemmältä ajalta, ja todellisuudessa liittymäpisterekisteri kattaa lähes kaikki kantakaupungin liittymät.

Liittymäpisterekisteriin on tätä tutkimusta varten viety liikennelaskentojen tietokannasta tieto jokaisen liittymän haaran liikennemäärästä sekä liikennelaskentavuodesta. Tarvittaessa puuttellisia liikennemäärätietoja on täydennetty liikennemallin avulla.

4.2 Aineiston raja

4.2.1 Käytettävän aineiston valinta

Työssä tarkastellaan vain liittymien liikenneturvallisuutta, minkä vuoksi työssä huomioidut onnettomuudet on rajoitettu vain liittymissä tai niiden välittömässä vaikutuspiirissä tapahtuneisiin onnettomuuksiin. Esimerkiksi katuosuuksilla, pihoilla tai pysäköintialueilla tapahtuneet onnettomuudet on rajattu pois aineistosta. Liittymäonnettomuudeksi on määritelty sellainen tapahtuma, jonka tapahtumapaikaksi onnettomuusrekisteriin on merkitty varsinainen liittymäalue.

Työssä käytetään vain valmiiksi saatavilla olevaa, koneellisesti käsiteltävää aineistoa. Onnettomuushavaintoja on aineistossa yli 30 000 ja liittymiä yli 3000, joten niiden tarkastaminen yksilöllisesti ei ole mahdollista. Jotta menetelmä soveltuu jatkuvaan, vuosittain toistuvaan käyttöön, tulee aineiston käsittelyn olla mahdollisimman pitkälle automatisoitua. Tästä syystä tässä työssä ei ole kerätty lisääaineistoa esimerkiksi maastokäynneillä.

4.2.2 Alueellinen ja ajallinen raja

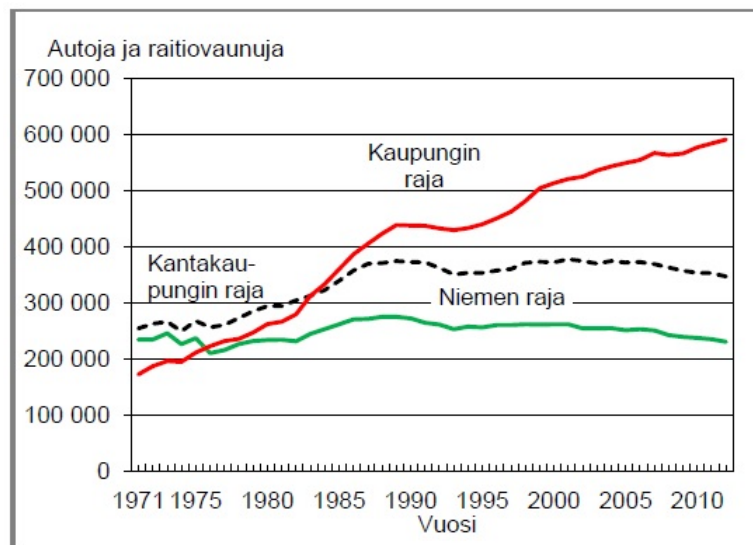
Tarkasteltujen liittymien joukko on rajattu maantieteellisesti kantakaupungin alueeseen. Tällä alueella liikenneverkko sekä liittymät ovat keskenään homogeenisempiä kuin muualla kaupungissa, mikä helpottaa mallintamista ja tulosten arviointia.

Esikaupunkialueille ovat ominaisia lukuisat pienet tonttikatujen liittymät, jotka vaikuttaisivat kaikkien liittymien onnettomuuksien odotusarvoon laskevasti. Näiden pienten liittymien liikennemäärätiedot eivät myöskään ole yhtä luotettavia kuin pääkatujen liittymien. Tonttikatujen liittymissä nopeudet ovat myös pääkatuja alhaisempia, mikä vähentää onnettomuuksien vakavuutta ja nostaa kynnystä ilmoittaa tapauksesta poliisille. Tämänkaltaiset

lievät onnettomuudet eivät siis usein päädy tässä työssä käytettävään onnettomuustilastoon.

Tässä työssä tarkasteltujakso on rajattu vuosiin 2005-2011. Merkittävin peruste rajaukselle on Helsingin kantakaupungin alueen nopeusrajoituksen aleneminen vuonna 2004. Nopeusrajoitus muuttui 40 kilometristä 30 kilometriin tunnissa, joten vuonna 2004 ja sitä ennen tapahtuneet onnettomuudet on jätetty aineistosta pois yhdenmukaisuuden vuoksi.

Liittymissä on käytettävissä vain yksi liikennemäärätieto kuvaamaan kaikkien havainnointivuosien liikennemäärää. Kuvasta 4.1 on havaittavissa, että kantakaupungin rajan ylittävä liikenne ei ole muuttunut merkittävästi 2000-luvulla. Samaan aikaan kuitenkin kaupungin rajan ylittävä liikenne on kasvanut lähes 20 %. Aineiston rajaaminen kantakaupunkiin pienentää liikennemäärätiedon puutteellisuudesta aiheutuvaa virhettä.



Kuva 4.1: Helsingin liikennemäärät (Helsingin kaupunki 2012)

4.3 Tutkimusaineiston käsittely

Mallintamista varten aineisto on muutettu taulukkomuotoon siten, että yksi rivi taulukossa kuvaa yhtä liittymää, ja yksi sarake yhtä mallissa käytettävää selittävää muuttujaa.

Onnettomuudet on laskettu onnettomuusaineistosta ja kirjattu kullekin liittymälle vuosikohtaisesti. Lopputuloksena tietokantataulukon yhdellä rivillä on yhden liittymän yhden vuoden onnettomuuksien lukumäärä. Tällöin seitsemän vuoden havaintojaksolla jokainen liittymä toistuu taulukossa seitsemän kertaa.

Työstettyyn aineistoon on luotu liittymille eri muuttujia alkuperäisen aineiston pohjalta. Muuttujien luomisen ja virheellisten liittymien poistamisen jälkeen aineistossa on 3 320 yksittäistä liittymää, eli 23 240 liittymä-vuosi-paria. Liittymät on lisäksi rajattu alueellisesti Helsingin kantakaupunkiin, ja rajauksen jälkeen aineisto käsittää 981 liittymää, eli 6 867 liittymä-vuosi-paria

4.3.1 Aineistosta luodut muuttujat

Alla on kuvattu mallintamisessa käytettyjen muuttujien luominen tutkimusaineiston pohjalta.

Liittymän liikennemäärä

Liittymäpisterekisterin aineistossa liikennemäärät on tilastoitu liittymille haarakohtaisesti. Aineistossa haarojen lukumäärä vaihtelee osalla liittymistä, ja osassa tapauksista sama haara toistuu tietokannassa useammin kuin kerran. Kaikki sellaiset liittymät, joilla on vähemmän kuin 3 tai enemmän kuin 4 haaraa tai joille on merkitty useita päällekkäisiä liikennemäärätietoja on poistettu aineistosta. Pääsuunta ja sivusuunta on määriteltä laskemalla keskiarvot vastakkaisten haarojen liikennemääräistä (1 ja 3 sekä 2 ja 4): pääsuunta on

haarapareista se jolla on enemmän liikennettä.

Suuntien määrittely ei perustu liittymien todellisiin liikennevirtojen suuntiin, vaan pelkkiin liikennemäärätietoihin. On olemassa esimerkiksi liittymiä, joissa kääntyvä suunta on liikenteen pääsuunta. Työssä on kuitenkin katsottu, että määritelmän tarkkuus on riittävä, eikä luotettavampaa tietoa liikenteen pääsuunnasta ole saatavissa.

Toinen työssä tarkasteltu tapa liikennemäärien huomioimiseksi on käyttää muuttujina liittymän koko liikennemäärää sekä sivusuunnan osuutta kokonaisliikennemäärästä. Eri liikennemäärämuuttujien vaikutus mallien hyvyyteen arvioidaan mallintamisen yhteydessä.

Nopeusrajoitus

Nopeusrajoitus on alkuperäisessä liittymäpisteaineistossa tilastoitu liittymille haarakohtaisesti. Liittymän nopeusrajoitukseksi on tässä työssä määritetty eri haaroista korkein nopeusrajoitus. Mikäli nopeusrajoitus on aineistossa jonkin haaran kohdalla 0 tai tyhjä arvo, on liittymä hylätty aineistosta.

Muuttujien välisiä korrelaatioita laskettaessa nopeusrajoitusta on käsitelty jatkuvana muuttujana. Mallin rakenteessa nopeusrajoitus ei kuitenkaan ole jatkuva, vaan nominaaliasteikollinen muuttuja. Jokainen liittymä kuuluu tiettyyn nopeusrajoitusluokkaan (esimerkiksi 30, 40 ja 50 km/h), joista jokaiselle on määriteltävissä mallissa luokan ominainen onnettomuustaso.

Liittymän haarojen lukumäärä

Liittymän haarojen lukumääräksi on määritelty 4 siinä tapauksessa, jossa haaran numero 4 liikennemäärä on positiivinen luku. Mikäli neljännen haaran liikennemäärä on 0 tai tyhjä muuttuja, on kyseessä kolmihaarainen liittymä. Sellaisissa tilanteissa joissa haarojen 1, 2 tai 3 liikennemäärä on 0 tai tyhjä muuttuja, liittymä on poistettu aineistosta. Haarojen lukumäärää kuvaava muuttuja on dikotominen, ja se saa arvon 0 silloin kun liittymä on kolmihaarainen, ja arvon 1 silloin kun liittymä on neliharainen.

Liittymän liikennevalo-ohjaus

Alkuperäisessä onnetomuustilastossa on merkintä tapahtumapaikan liikennevalopisteen numerosta. Liittymä on merkitty liikennevalo-ohjauksiseksi, mikäli liittymässä on tapahtunut onnettomuus, jossa liikennevalopisteen numero on jokin muu kuin puuttuva luku. Liikennevalo-ohjausta kuvaava muuttuja saa arvon 0 kun liittymä on valo-ohjaukseton ja arvon 1 silloin kun kyseessä on liikennevalo-ohjauksinen liittymä.

Määritelmässä on periaattellinen ongelma: liikennevalo-ohjauksinen liittymä jossa ei ole koskaan tapahtunut onnettomuutta ei voi saada merkintää liikennevalo-ohjauksesta mallintamisaineistoon, sillä liittymä ei tällöin esiinny onnettomuusrekisterissä lainkaan. Käytetty liittymäpisterekisteri sisältää vain sellaiset liittymät joissa on tapahtunut tilastoitu onnettomuus, mutta onnettomuuksien tarkasteltu ajanjakso on pidempi kuin onnettomuusaineistossa. Ongelma koskee siis sellaisia liittymiä, joissa on liikennevalo-ohjaus, mutta joissa ei ole tapahtunut yhtään poliisin tietoon tullutta onnettomuutta vuoden 2003 jälkeen. Tämän työn tapauksessa ongelman kuitenkin oletetaan koskevan niin pientä määrää liittymiä, ettei se vaikuta onnettomuusmallien tarkkuuteen.

4.3.2 Hylätyt muuttujat

Työssä tarkasteltiin myös muita muuttujia, joita ei kuitenkaan käytetty varsinaisessa mallintamisessa.

Katuluokka

Katuluokka on merkitty liittymäpisterekisteriin jokaiselle liittymän haaralle. Katuluokan kuitenkin oletetaan kuvaavan liikenteen kannalta samoja toiminnallisia tekijöitä kuin tiedot liikennevaloista, nopeusrajoituksesta ja liikennemääristä. Katuluokan tapauksessa on myös vaikeaa tulkita muuttujan periaatteellista vaikutusta liikenneonnettomuuksiin sellaisessa tilantees-

sa jossa liittymän eri haarojen katuluokat eroavat toisistaan. Mallintamises-
sa pelkkä muuttujan aiheuttama parannus mallin hyvydessä ei ole riittävä,
vaan muuttujan tulee olla myös loogisesti perusteltavissa.

Olosuhteisiin liittyvät muuttujat

Onnettomuusaineistossa on tiedot useimpien onnettomuuksien tapahtumiso-
losuhteista, kuten säätilasta, onnettomuuden osapuolten veren alkoholipitoi-
suudesta sekä siitä olivatko liikennevalot toiminnassa. Nämä tekijät eivät
kuitenkaan ole tämän työn kannalta relevantteja, sillä ne eivät kuvaa liit-
tymän ominaisuuksia, vaikka saattavatkin selittää yksittäisen onnettomuu-
den syytä.

4.4 Muuttujien arviointi

4.4.1 Muuttujien korrelaatiot

Taulukossa 4.1 on esitetty edellä määriteltyjen muuttujien keskinäiset korrelaatiot, eli muuttujien korrelaatiomatriisi. Korrelaatiokertoimenä on käytetty muuttujien välistä lineaarista riippuvuutta kuvaavaa Pearsonin korrelaatiokerrointa, joka saa arvoja väliltä $-1...1$.

Taulukko 4.1: Kantakaupungin kaikkien liittymien korrelaatiomatriisi

	Pääsuunta	Sivusuunta	Nelihaara	Liikennevalot	Nopeusraaj.
Pääsuunta	1.00	0.65	0.00	0.55	0.76
Sivusuunta		1.00	-0.05	0.50	0.54
Nelihaara			1.00	0.21	-0.02
Liikennevalot				1.00	0.45
Nopeusraaj.					1.00

Tuloksista on havaittavissa selvä positiivinen korrelaatio liittymän pää- ja sivusuuntien liikennemäärien välillä. Liikennemäärien ja liittymän haarojen lukumäärien välillä toisaalta ei ole lainkaan korrelaatiota. Voimakkain korrelaatio on nopeusrajoituksen ja liikennemäärien välillä. Nopeusrajoitukseltaan korkeampien liittymien liikennemäärä on siis keskimäärin suurempi kuin nopeusrajoitukseltaan alhaisten liittymien.

Haarojen lukumäärä ei korreloi merkittävästi yhdenkään muun muuttujan kanssa, eli se soveltuu hyvin regressiomallin selittäväksi muuttujaksi. Liikennevalo-ohjauksella puolestaan on selvä positiivinen korrelaatio kaikkien muiden muuttujien kanssa. Mallintamisessa multikollineaarisuudesta johtuvia ongelmia on tällöin mahdollista lieventää jakamalla liittymät kahteen eri kategoriaan: liikennevalo-ohjauksisiin ja valo-ohjauksettomiin.

Taulukoissa 4.2 sekä 4.3 on esitetty korrelaatiomatriisit kummallekin eri liit-

tymäkategorialle.

Taulukko 4.2: Kantakaupungin valo-ohjauksettomien liittymien korrelaatiomatriisi

	Pääsuunta	Sivusuunta	Nelihaara	Nopeus
Pääsuunta	1.00	0.60	−0.07	0.74
Sivusuunta		1.00	−0.08	0.50
Nelihaara			1.00	−0.12
Nopeus				1.00

Taulukko 4.3: Kantakaupungin liikennevalo-ohjauksisten liittymien korrelaatiomatriisi

	Pääsuunta	Sivusuunta	Nelihaara	Nopeus
Pääsuunta	1.00	0.47	−0.29	0.58
Sivusuunta		1.00	−0.35	0.36
Nelihaara			1.00	−0.16
Nopeus				1.00

Liikennevalo-ohjauksissa liittymissä korrelaatio haarojen lukumäärän ja muiden muuttujien välillä on selvästi voimakkaampi kuin valo-ohjauksettomissa liittymissä.

Nopeusrajoitus kuitenkin korreloi vähemmän liikennemäärien kanssa liikennevalo-ohjauksissa liittymissä kuin valo-ohjauksettomissa liittymissä.

Luku 5

Tutkimusprosessi ja tutkimuksen tulokset

5.1 Onnettomuusmallien laatiminen

5.1.1 Mallintamisprosessin kuvaus

Empirical Bayes -menetelmää varten työssä on laadittu regressiomalli, jolla pyritään ennustamaan yksittäisen liittymän onnettomuusfrekvenssiä. Mallintamisprosessin tavoitteena on ollut laatia kaksi eri mallia: toinen kantakaupungin liikennevalo-ohjauksisille liittymille ja toinen kantakaupungin valo-ohjauksettomille liittymille. Liittymien jako kahteen eri kategoriaan mallintamista varten on perustelu aineistoa käsittelevässä luvussa.

Eri mallit on laadittu siten, että jokaiseen malliin on lisätty yksi uusi selittävä muuttuja kerrallaan. Käytettyjä selittäviä muuttujia ovat pääsuunnan liikennemäärä, sivusuunnan liikennemäärä, liittymän haarojen lukumäärä sekä nopeusrajoitus. Muuttujien lisäksi malleissa on mukana myös muuttujien yhteisvaikutus, *interaktiotermi*, eli esimerkiksi se että liittymä on nelihaarainen ja sen nopeusrajoitus on 40 kilometriä tunnissa. Yksinkertainen malli on ns.

nollamalli, eli malli jossa on vain vakiotermi.

Taulukossa 5.1 on esitetty mallien rakenteet ja numerointi, joilla eri malleihin viitataan jatkossa. Merkintä ”*” muuttujien välissä tarkoittaa sitä, että myös muuttujien yhteisvaikutus on mallissa mukana. Kaikki työssä tarkastellut mallit on esitetty yksityiskohtaisesti liitteessä A.

Taulukko 5.1: Työssä tarkasteltujen mallien rakenteet

Malli	Muuttujat
0	Pelkkä vakiotermi
1	Pääsuunta
2	Pääsuunta + Sivusuunta
3	Pääsuunta + Sivusuunta + Nelihaara
4	Pääsuunta + Sivusuunta + Nelihaara + Nopeus
5	Pääsuunta *Sivusuunta + Nelihaara + Nopeus
6	Pääsuunta *Sivusuunta *Nelihaara + Nopeus
7	Pääsuunta *Sivusuunta *Nelihaara *Nopeus

Taulukoissa 5.2 sekä 5.3 on esitetty eri mallien Akaiken informaatiokriteerit (AIC), sekä muutokset mallin Akaiken informaatiokriteerissä kun malliin lisätään selittäviä muuttujia. Sarake ”Muutos” on kahden mallin välinen erotus Akaiken informaatiokriteerin arvossa, ja sarake ”Kumul.muutos” mallin kumulatiivinen muutos verrattuna nollamalliin.

Burnhamin ja Andersonin (2002) mukaan jo suuruusluokkaa 2 oleva ero on tilastollisesti merkitsevä kaikilla yleisillä merkitsevyystasoilla, ja on merkki toisen mallin paremmuudesta. Mitä pienempi mallin AIC on, sitä paremmin malli sopii aineistoon.

Taulukko 5.2: Kantakaupungin valo-ohjauksettomien liittymien onnettomuusmallit

Malli	AIC	Muutos	Kumul.muutos
0	11644		0
1	11218	−426	−426
2	11162	−56	−482
3	10914	−248	−730
4	10878	−36	−766
5	10658	−220	−985
6	10662	4	−981
7	10464	−199	−1180

Taulukko 5.3: Kantakaupungin liikennevalo-ohjauksisten liittymien onnettomuusmallit

Malli	AIC	Muutos	Kumul.muutos
0	7545		0
1	7521	−24	−24
2	7477	−44	−68
3	7344	−134	−201
4	7238	−105	−306
5	7190	−48	−355
6	7152	−38	−393
7	7083	−69	−462

Liikennevalo-ohjauksissa liittymissä AIC:n kehitys on johdonmukainen: jokaisen uuden muuttujan lisääminen parantaa mallin sopivuutta aineistoon. Muutos kuitenkin pienenee siinä vaiheessa kun lisätermit ovat muuttujien yhteisvaikutusta kuvaavia kertoimia, eli siirryttäessä mallista 4 malliin 5.

Valo-ohjauksettomissa liittymissä puolestaan muutos on epäsäännöllisempi, ja mallissa 5 mukana oleva pää- ja sivusuunnan liikennemäärien yhteisvaikutus parantaa mallin sopivuutta huomattavan paljon. Liikennemäärät, haarojen lukumäärä sekä nopeusrajoitus ovat molemmissa liittymätyypeissä selvästi mallin selitysastetta parantavia muuttujia.

Valo-ohjauksettomissa liittymissä nopeusrajoitukset 40, 50 ja 70 km/h eivät ole tilastollisesti merkitseviä regressiokertoimia (kts. taulukko liitetaulukko A.4). Tästä syystä valo-ohjauksettomien liittymien onnettomuusmallissa nopeusrajoitusta ei voida käyttää selittävänä muuttujana. Johdonmukaisuuden vuoksi nopeusrajoitusta ei käytetä tämän vuoksi myöskään liikennevalo-ohjauksisten liittymien mallissa.

Muuttujien yhteisvaikutuksen kanssa kaikki yhteistermit eivät ole tilastollisesti merkitseviä. Tämä tarkoittaisi Empirical Bayes –menetelmän kannalta sitä, että EB-estimaattia laskettaessa huomioitaisiin esimerkiksi yhteisvaikutus siitä, että liittymä on nelihaarainen ja nopeusrajoitus on 40 km/h, mutta ei kuitenkaan sitä että liittymä on nelihaarainen ja nopeusrajoitus on 50 km/h. Tämä lähestymistapa ei olisi loogisesti perusteltavissa, ja se lisäksi monimutkaistaisi mallin käyttämistä ja uudelleenestimointia tulevaisuudessa.

Mallin yksinkertaistamisen vuoksi tässä työssä käytettäväksi mallin rakenteeksi valitaan se, jossa muuttujina ovat pääsuunta, sivusuunta, sekä haarojen lkm. Taulukon 5.1 notaation mukaisesti tämä on malli numero kolme.

5.1.2 Liikennemäärämuuttujan arviointi

Työssä on verrattu kahta eri vaihtoehtoa liikennemäärien huomioimiseksi mallissa. Ensimmäisessä vaihtoehdossa, josta käytetään nimitystä tyyppi A, on muuttujina pääsuunnan liikennemäärä sekä sivusuunnan liikennemäärä. Toisessa vaihtoehdossa, josta käytetään nimitystä tyyppi B, muuttujina ovat koko liittymän liikennemäärä sekä sivusuunnan liikennemäärän suhteellinen

osuus kokonaisliikennemäärästä.

Taulukossa 5.4 on vertailtu Akaiken informaatiokriteerin arvoja tyyppien A ja B mallien välillä. Mallien numerointi on sama kuin taulukossa 5.1 on esitetty. Esimerkiksi mallissa numero 2 muuttujina on A-tyypin tapauksessa pääsuunnan liikennemäärä sekä sivusuunnan liikennemäärä, ja B-tyypin tapauksessa liittymän kokonaisliikennemäärä sekä sivusuunnan osuus kokonaisliikennemäärästä. Erotus-sarakkeeseen on merkitty B-tyypin ja A-tyypin mallien Akaiken informaatiokriteerien erotukset, ja positiivinen erotus tarkoittaa A-tyypin mallin olevan parempi kyseiselle muuttujakombinaatiolle.

Taulukko 5.4: AIC:n arvot eri liikennemäärämuuttujilla estimoiduille malleille

Malli	Tyyppi A	Tyyppi B	Erotus
Valo-ohjauksettomat liittymät			
1	11218	11166	-52
2	11162	11167	5
3	10914	10918	4
4	10878	10880	3
5	10658	10879	221
6	10662	10872	210
7	10464	10659	195
Liikennevalo-ohjauksiset liittymät			
1	7521	7493	-28
2	7477	7457	-20
3	7344	7328	-16
4	7238	7228	-11
5	7190	7225	34
6	7152	7170	18
7	7083	7125	42

Valo-ohjauksettomissa liittymissä A-tyypin malli tarjoaa paremman sovituksen havaintoaineistoon, ja Liikennevalo-ohjauksisissa liittymissä B-tyypin malli on parempi mallivaihtoehdoille 1-4. Kummassakin kategoriassa B-tyyppi on selvästi huonompi kun muuttujien yhteisvaikutukset otetaan mukaan tarkasteluun.

Työn kannalta merkityksellisin on A- ja B-tyyppien ero luokan 3 malleissa, joka on valittu työssä käytettäväksi. B-tyypin ero A-tyyppiin nähden

liikennevalo-ohjauksisissa liittymissä on suurempi kuin A-tyyppin ero B-tyyppiin valo-ohjauksettomissa liittymissä, mutta kumpikin erotus on tilastollisesti merkitsevä.

Kummankin mallivaihtoehdon valinta on perusteltua tässä työssä. A-tyyppin mallille on kuitenkin enemmän tukea työssä käytetyssä kirjallisuudessa: tarkastelluissa tutkimuksissa käytettyjen onnettomuusmallien muuttujina ovat tyypillisesti eri suuntien liikennemäärät. Pää- ja sivusuunnat on lisäksi määriteltä työssä liikennemäärien perusteella, eivätkä ne perustu esimerkiksi liittymien liikennesuunnitelmiin. Tämä, sekä liikennemäärätietojen epätarkkuus vaikeuttavat liikennemäärien suhdelukujen laskentaa. Riskinä on, että suhdelukujen laskennassa liikennemäärien ja suuntien määrittelyn mahdolliset epätarkkuudet kertaantuvat.

Tällä perusteella työssä valitaan käytettäväksi liikennemäärämuuttujiksi pää- sekä sivusuuntien liikennemäärät.

5.2 Valitut onnettomuusmallit

5.2.1 Mallien esittely

Työssä käytettäviä malleja on kaksi: liikennevalo-ohjauksisten ja valo-ohjauksettomien liittymien mallit. Mallien selittävinä muuttujina ovat pääsuunnan liikennemäärä, sivusuunnan liikennemäärä sekä liittymän haarojen lukumäärä. Onnettomuuksien odotusarvo yksittäiselle liittymälle on tällöin:

$$\mu = e^{\alpha + \beta_1 x_{Paasuunta} + \beta_2 x_{Sivusuunta} + \beta_3 x_{Nelihaara}} \quad (5.1)$$

Taulukoissa 5.5 ja 5.6 on esitelty mallien regressiokertoimet sekä niiden arvot. Yksityiskohtaiset tiedot malleista, kuten regressiokertoimien luottamusvälit ja p-arvot on esitetty liitteessä A.

Taulukko 5.5: Kantakaupungin valo-ohjauksettomien liittymien onnettomuusmalli

Vakiotermi	Paasuunta	Sivusuunta	Nelihaara	Hajontaparametri
-1.100854	0.000038	0.000063	0.736647	0.753624

Taulukko 5.6: Kantakaupungin liikennevalo-ohjauksisten liittymien onnettomuusmalli

Vakiotermi	Paasuunta	Sivusuunta	Nelihaara	Hajontaparametri
0.349682	0.000008	0.000032	0.534167	0.377665

5.2.2 Mallien arviointi

Mallien tarkkuutta on tämän työn tapauksessa mahdollista arvioida tutkimalla niiden hajontaparametrin vaikutusta Empirical Bayes -luvun muodostamiseen. EB-luku on painotettu keskiarvo mallin ennustamasta onnettomuusfrekvenssin odotusarvosta sekä liittymässä havaitusta onnettomuusfrekvenssistä, jossa mallin tuottama odotusarvo saa painokertoimen w siten, että

$$w = \frac{1}{1 + \frac{E(Y)}{k}}$$

$E(Y)$ on mallin perusteella määritelty odotusarvo liittymän onnettomuusfrekvenssille ja k on mallin hajontaparametrin käänteisluku.

Taulukossa 5.7 on havainnollistettu Empirical Bayes -luvun laskemisessa käytetyn painokertoimen arvoa eri mallityypeille ja eri mallin tuottamille odotusar-

voille. Taulukosta on havaittavissa tapa jolla menetelmä huomioi luvussa 2.1 esitellyn *regression keskiarvoa kohti*: mallin odotusarvon lähestyessä nollaa havaitun onnettomuusfrekvenssin painoarvo estimaatissa laskee. Tämä on selitettävissä oletuksella siitä, että havaittu onnettomuusfrekvenssi joka on suhteettoman suuri verrattuna odotusarvoon johtuu todennäköisesti satunnaisvaihtelusta.

Taulukko 5.7: Painokertoimen arvot eri odotusarvoille

Tyyppi	1	2	3	4	5
Valo-ohjaukseton	0, 57	0, 40	0, 31	0, 25	0, 21
Valo-ohjauksinen	0, 73	0, 57	0, 47	0, 40	0, 35

5.2.3 Regressiokertoimien tulkinta

Mallin tarkoituksena on ennustaa yksittäisen liittymän onnettomuusfrekvenssi sen tunnettujen parametrien funktiona. Mallien regressiokertoimien yksittäisillä lukuarvoilla ei ole merkitystä, eikä niitä tule arvioida ilman mallin kontekstia. Regressiokertoimet kuvaavat vain sitä kuinka eri tekijät vaikuttavat onnettomuuksien lukumäärään tässä työssä käytetystä aineistossa estimoidussa mallissa.

Esimerkiksi liittymän haarojen lukumäärää kuvaavan regressiokertoimen lukuarvosta ei voida päätellä suoraan liittymän nelihaaraisuuden vaikutusta liittymän liikenneturvallisuuteen. Luvussa 3.3.1 kuvatun multikollineaarisuuden vuoksi regressiokertoimien vaikutusta ennustettavan muuttujan arvoihin ei voida päätellä ilman tietoa muista muuttujista. Nelihaaraisuuden vaikutuksen tutkiminen edellyttäisi tarkoitusta varten tehtyä koejärjestelyä ja muiden tekijöiden kontrollointia.

5.3 Mustat pisteet

5.3.1 Empirical Bayes -menetelmällä tunnistetut kohteet

Taulukossa 5.8 on esitetty 20 kantakaupungin liittymää, joilla on vuonna 2011 ollut suurin EB-estimaatti. Taulukossa on myös esitetty liittymäkohtaisesti PFI-luku, mallin tuottama ennuste onnettomuusfrekvenssille ($E(Y)$), liittymässä vuonna 2011 havaittu onnettomuuksien lukumäärä (y_i) sekä liittymän sijoitus IND5-luvun perusteella, mikäli liittymä on kuulunut 20 vaarallisimman liittymän joukkoon Helsingin kaupungin (2012) julkaisemassa onnettomuustilastossa. Taulukkoon on lisäksi merkitty lihavoituina ne kohteet, jotka eivät lainkaan olleet mukana Helsingin kaupungin onnettomuustilastossa, eli ne liittymät jotka eivät IND5-luvun perusteella ole mustia pisteitä.

Taulukko 5.8: Kantakaupungin 20 vaarallisinta liittymää vuonna 2011

	Ruutupiste	Osoite	EB	PFI	$E(Y)$	y_i	IND5 sij.
1	G3T215	Mannerh.t. - Kaivok. - Simonk.	14.15	9.74	4.41	20	3
2	H4P236	Mäkelänk. - Sturenk.	10.92	4.80	6.12	13	8
3	G3T137	Mannerh.t. - Pohjoisesp. - Lönnrotink.	10.46	6.61	3.85	15	6
4	G3S412	Mannerh.t. - Helsingink. - Runebergink.	9.39	3.55	5.84	11	1
5	G3S128	Mechelinink. - Lapinlahdent.	8.89	4.04	4.85	10	—
6	G4P407	Mannerh.t. - Koroistent. - Kytösuont.	8.52	4.84	3.68	12	16
7	G4T115	Helsingink. - Sturenk. - Kolmas linja	8.23	3.30	4.93	10	—
8	G3S341	Mannerh.t. - Töölönlahdenk.	8.02	5.56	2.46	14	5
9	G4R208	Mannerh.t. - Reijolank. - Stenbäckink.	7.91	3.49	4.42	10	—
10	G3S139	Mechelinink. - Pohj. Rautatiek.	7.20	4.94	2.26	13	2
11	G3T111	Lönnrotink. - Annank.	6.70	3.72	2.98	10	—
12	G4S203	Nordenskiöldink. - Veturit.	6.63	3.13	3.50	9	—
13	G4T313	Ratapihant. - Asemapäällikönk.	6.53	2.33	4.20	8	—
14	G4R203	Topeliuksenk. - Tukholmank.	6.51	2.34	4.17	8	10
15	F5T107	Vihdintie - Lapinmäent.	6.40	2.41	3.99	8	—
16	G4P309	Mannerh.t. - Kiskontie	6.28	4.29	1.99	12	—
17	G4R130	Nordenskiöldink. - Urheiluk.	5.97	2.59	3.38	8	—
18	H3P301	Unionink. - Liisank. - Kaisaniemenk.	5.83	1.79	4.04	7	—
19	G3S409	Mannerh.t. - Kivelänk.	5.81	2.75	3.06	7	—
20	G3S221	Fredrikink. - Pohj. Rautatiek.	5.74	2.64	3.10	8	13

Esitetty taulukko on vain ote työssä tehdystä analyysistä. Kaikkien aineistossa mukana olleiden liittymien liikenneturvallisuus on määritetty jokaiselle tarkasteluvuodelle, ja mallintamisaineisto luovutetaan työn liitteenä.

5.3.2 Empirical Bayes -menetelmän arviointi

Empirical Bayes -luvun avulla tunnistettujen mustien pisteiden joukko on johdonmukainen ja uskottava. Työn aikana käytyjen keskustelujen perusteella vaarallisimmat kohteet ovat yhteneviä suunnittelijoiden arkikokemuksen ja liittymien luonteen kanssa.

Empirical Bayes -menetelmän ero aiempaan IND5-lukuun on havaittavissa liittymistä, jotka on merkitty lihavoidulla tekstillä taulukkoon 5.8. Risteyksissä Mechelininkatu - Lapinlahdentie, Helsinginkatu - Sturenkatu - Kolmas linja sekä Lönnrotinkatu - Annankatu on tapahtunut vuonna 2011 vähintään kymmenen liikenneonnettomuutta, mutta niistä yksikään ei ole johtanut henkilövahinkoihin. Koska IND5-luvussa omaisuusvahinko-onnettomuuksien painoarvo henkilövahinko-onnettomuuksiin nähden on vain viidesosa, eivät nämä liittymät ole IND5-luvun perusteella mustia pisteitä.

Osassa liittymistä, kuten Mechelininkadun ja Lapinlahdenkadun liittymässä on kuitenkin perusteltua olettaa, että monet tapahtuneista onnettomuuksista olisivat voineet johtaa henkilövahinkoihin. Samankaltaisia liittymiä verrattaessa henkilövahinkojen puuttuminen ole välttämättä osoitus liittymän paremmasta turvallisuudesta. Tämä on kenties merkittävin periaattellinen ero nykyisin käytetyn IND5-menetelmän ja Empirical Bayes -menetelmän välillä.

Taulukosta 5.8 on myös havaittavissa, että Helsinginkadun ja Sturenkadun liittymän sekä Mannerheimintien ja Töölönlahdenkadun liittymän liikenneturvallisuus on EB-luvun perusteella lähes sama. PFI-luku on kuitenkin Mannerheimintie - Töölönlahdenkatu -liittymälle suurempi, eli kyseisen liittymän liikenneturvallisuuden poikkeama liittymältä odotetusta tasosta on suurempi

kuin Sturenkadun ja Helsinginkadun liittymällä.

5.4 IND5- ja EB-menetelmien vertailu

Taulukossa 5.9 on esitetty luvussa 3.4 kuvatun konsistenssitestin tulos. Luvut kertovat kuinka monta kertaa vuoden 2011 20 vaarallisinta liittymää esiintyvät sarakkeen mukaisen vuoden 20 vaarallisimman liittymän joukossa.

Taulukko 5.9: Menetelmien konsistenssi vertailuvuosina

Menetelmä	2011	2010	2009
IND5	20	10	12
EB	20	13	10
PFI	20	11	7

Taulukossa 5.10 on esitetty sellaisten liittymien lukumäärä, jotka toistuvat kunakin vuonna sekä EB- että PFI-menetelmän tuottamassa 20 vaarallisimman kohteen listassa.

Taulukko 5.10: EB- ja PFI-menetelmien yhteneväisyys eri vuosina

	2011	2010	2009
EB ja PFI	18	17	3

Vertailun perusteella IND5 ja EB eivät eroa tämän työn tapauksessa toisistaan havaittavalla tavalla. PFI on käytettyjen testien perusteella selvästi kahta muuta mittaria epäkonsistentimpi eri vuosien välillä. Nämä tulokset ovat yhteneviä tarkastellun kirjallisuuden kanssa. Yksinkertainen onnettomuusfrekvenssiin perustuva menetelmä on esimerkiksi Montellan (2010) tutkimuksessa osoittautunut heti EB-menetelmän jälkeen parhaaksi. Onnetto-

muusmallien hajonnan kasvaessa havaitun onnettomuusfrekvenssin painotus kasvaa EB-menetelmässä, jolloin menetelmät lähestyvät toisiaan.

Vähäinen ero menetelmien välillä selittyy myös osittain testin alkeellisuudella. Liittymien järjestysluku ei ole hyvä mittari menetelmien arviointiin, sillä tulosten perusteella eri liittymien lukuarvot ovat sekä IND5- että EB-luvuissa hyvin lähellä toisiaan. Esimerkiksi taulukossa 5.8 sijoituksilla 11-20 olevien liittymien EB-lukujen ero on vähemmän kuin 1 onnettomuus vuodessa. Kumpikaan verratuista menetelmistä ei kykene täysin poistamaan satunnaisvaihtelun vaikutusta, minkä vuoksi pienten erojen arviointi on epätarkkaa.

Luku 6

Yhteenveto ja johtopäätökset

6.1 Yhteenveto

Tässä diplomityössä on sovellettu rajatussa mittakaavassa uutta menetelmää liittymien liikenneturvallisuuden mittaamiseksi Helsingissä.

Työn teoreettisessa osassa on kuvattu liikenneturvallisuuden määritelmä sekä perusteltu liikenneturvallisuuden mittaaminen onnettomuusfrekvenssin avulla. Työssä on myös tehty kirjallisuuskatsaus aiempaan tutkimukseen vaarallisten kohteiden tunnistamismenetelmistä ja perusteltu työssä käytettävän Empirical Bayes -menetelmän valinta.

Työn empiirisessä osassa on sovellettu valittua Empirical Bayes -menetelmää liikenneturvallisuuden mittaamiseksi Helsingin kantakaupungin liittymissä. Käytetty menetelmä mittaa liikenneturvallisuutta havaitun onnettomuusfrekvenssin sekä liittymältä odotetun onnettomuusfrekvenssin painotettuna keskiarvona. Odotettu onnettomuuksien lukumäärä määritellään tilastollisen mallintamisen avulla, ja menetelmän painokerroin määräytyy mallin tarkkuuden perusteella.

Työssä on laadittu Empirical Bayes -menetelmän edellyttämä onnettomuusmalli, joka ennustaa liittymässä vuosittain tapahtuvien onnettomuuksien lu-

kumäärää liittymän tunnettujen parametrien funktiona. Onnettomuusmalli on tyypiltään negatiivinen binomimalli, joka kuuluu yleistettyjen lineaaristen mallien perheeseen.

Tutkimusaineistona on käytetty Helsingin kaupungin keräämää onnettomuusaineistoa sekä liittymäpisterekisteriä. Onnettomuusaineistossa on tiedot tapahtuneista onnettomuuksista, sisältäen muun muassa tapahtumapaikan, tapahtuma-ajan, osallisten tiedot sekä säätilan. Liittymäpisterekisteriin on tallennettu tiedot Helsingin liittymistä, kuten liittymän eri haarojen liikennemäärät sekä nopeusrajoitus.

Mallintamisessa selittävinä muuttujina on käytetty liittymän pää- ja sivusuuntien liikennemääriä sekä liittymän haarojen lukumäärää. Liikennevalo-ohjauksille ja valo-ohjauksettomille liittymille on laadittu omat mallinsa. Liittymän nopeusrajoitus hylättiin mallintamisvaiheessa selittävien muuttujien joukosta, sillä se ei osoittautunut tilastollisesti merkittäväksi kaikilla liittymätyypeillä.

Mallin perusteella jokaiselle liittymä-vuosi-parille on laskettu Empirical Bayes -luku sekä Potential for improvement -luku, jotka kuvaavat liittymän turvallisuutta.

Empirical Bayes -tunnusluku lasketaan kaavalla:

$$EB = wE(Y) + (1 - w)y \quad (6.1)$$

jossa $E(Y)$ on onnettomuusmallin liittymälle ennustama onnettomuusfrekvenssi ja y on havaittu onnettomuusfrekvenssi. Painokerroin w on

$$w = \frac{1}{1 + \frac{E(Y)}{k}} \quad (6.2)$$

jossa k on onnettomuusmallin hajontaparametrin käänteisluku.

PFI-luku lasketaan Empirical Bayes -luvun ja onnettomuusmallin ennusteen $E(Y)$ erotuksena.

Varsinaiset mustat pisteet on tunnistettu asettamalla liittymät järjestykseen EB-luvun, eli liittymän vaarallisuuden perusteella. Menetelmän avulla tunnistetut mustat pisteet ovat uskottavia ja johdonmukaisia, mutta työssä ei kyetty määrittämään selvää tilastollisesti merkitsevää eroa eri menetelmien suorituskyvyn välillä.

6.2 Johtopäätökset

Tämän diplomityön tavoitteena on ollut luoda edellytykset Empirical Bayes -menetelmän käyttöönotolle Helsingin kaupungin liikenneturvallisuustyössä. Helsingin kaupunkisuunnitteluviraston liittymistä ja onnettomuuksista ylläpitämä tilastoaineisto on osoittautunut mallintamiseen soveltuvaksi. Aineiston pohjalta on kyetty laatimaan negatiivinen binomimalli, jolla voidaan selittää liittymien onnettomuusfrekvenssiä liittymän eri ominaisuuksien avulla.

Empirical Bayes -menetelmän paremmuutta nykyisin käytettyyn menetelmään nähden työssä ei kyetty määrittämään laskennallisesti. Kirjallisuus kuitenkin tukee vahvasti Empirical Bayesin käyttöä. Lisäksi EB-luvun ohella laskettava PFI-luku helpottaa liittymän turvallisuuden arviointia suhteessa muihin vastaaviin liittymiin.

Työn tavoitteet ja menetelmät ovat olleet hyvin samankaltaisia työssä tarkasteltujen aiempien tutkimusten kanssa. Esimerkiksi Kulmalan (1995) väitöstutkimuksessa on toteutettu vastaava onnettomuusmallien laadinta sekä Empirical Bayes -lukujen määrittäminen maanteiden tasoliittymille. Katuverkon liittymillä vastaavaa tutkimusta ei kuitenkaan tiettävästi ole Suomessa aiemmin toteutettu. Kahdessa työssä tarkastellussa tutkimuksessa (Kusumawati & Wong 2010 sekä Miao & Lord 2005) on sovellettu Empirical Bayes -menetelmää kaupunkimaisen ympäristön liittymiin, mutta mallit eivät ole sellaisenaan vertailukelpoisia Helsingin, Toronton ja Singaporen välillä.

Menetelmää on sovellettu pilottihankkeen kaltaisesti vain kantakaupungin tasoliittymiin, ja koko kaupungin liittymien turvallisuuden arvioimiseksi me-

netelmä tulee laajentaa käsittämään myös muut Helsingin osa-alueet.

6.3 Ehdotukset jatkotoimenpiteiksi

6.3.1 Menetelmän laajentaminen muualle Helsinkiin

Tässä työssä laadittu menetelmä on sellaisen sovellettavissa vain kanta-kaupungin alueelle, sillä käytetty tilastollinen malli on estimoitu vain kanta-kaupungin liittymistä. Esikaupunkialueille luonteenomaista on pienten tonttikatujen liittymien suuri lukumäärä. Näiden liittymien liikennemääräksi on useassa tapauksessa merkitty liittymäpisterekisteriin 100, mikä on osoitus liikennemäärätiedon puutteellisuudesta.

Esikaupunkialueiden mallintaminen edellyttäisi nykyistä tarkempaa liikennemäärätietoa pienistä liittymistä sekä myös historiatietoa liikennemäärien kehityksestä. Jotta mallin parametrit voidaan estimoida oikein, tulee liittymäaineistossa kuvata jokaisen vuoden liikennemäärä kussakin liittymässä. Muualla kuin kantakaupungissa liikennemäärät ovat kasvaneet 2000-luvulla, joten malli on osittain virheellinen mikäli tätä kasvua ei huomioida. Tästä aiheutuvan virheen merkitystä menetelmän tarkkuudelle on kuitenkin vaikea arvoida.

6.3.2 Menetelmän laajentaminen eri liittymätyyppeihin

Tärkein laajennuskohde menetelmälle ovat kantakaupungin liittymät, jotka eivät ole kolmi- tai nelihaaraisia, tai jotka ovat muusta syystä poistettu pilot-tivaiheen mallintamisaineistosta. Tällaisia liittymiä ovat esimerkiksi Viiskulma sekä Nordenskiöldinaukio. Nordenskiöldinaukion kaltaista monimutkaista liittymäkokonaisuutta voidaan mallintamisen näkökulmasta kuvata sarjana yksittäisiä kolmi- ja nelihaaraliittymiä mikäli onnettomuudet kyetään paikantamaan riittävän tarkasti liittymäalueen sisällä. Vaihtoehtoisesti kaikkia

selvästi muista poikkeavia liittymiä voidaan kuvata yhdellä kategorialla.

Laaajoja eritasoliittymiä tai muita liittymäkokonaisuuksia ei ole huomioitu tässä työssä. Nämä ovat kuitenkin liikennemääriltään suurimpia kohteita Helsingissä, ja niiden sisällyttäminen mallintamiseen ja EB-estimaatteihin on jatkossa toivottavaa. Tämä edellyttää esimerkiksi moottoritien eritasoliittymän tapauksessa päätöstä siitä millä tavalla liittymä esitetään tilastoaineistossa. Yksi esimerkkikeino on kaikkien liittymän ramppien esittäminen yhtenä liittymäpisteenä siten, että liikenteen suuntia voi olla enemmän kuin kaksi. Toisena mahdollisena vaihtoehtona jokainen liittymän ramppi voidaan esittää omana liittymäpisteenään, jolloin selittävinä muuttujina ovat esimerkiksi rampista pääväylälle liittyvä liikennemäärä sekä koko väylän liikennemäärä.

Eritasoliittymien lukumäärä on kuitenkin melko pieni koko pääkaupunkiseudunkin alueella, ja liittymät ovat fyysisiltä ominaisuuksiltaan yksilöllisiä. Tämänkaltaisessa tilanteessa tilastollisen onnettomuusmallin laatiminen ei välttämättä ole perusteltua, vaan liittymien liikenneturvallisuutta voidaan arvioida pelkällä toteutuneiden onnettomuuksien asiantuntija-arviolla.

6.3.3 Liittymäkohtaisten tietojen tarkentaminen

Tieto jalankulkijoiden ja pyöräilijöiden lukumääristä saattaisivat parantaa mallien selityskykyä etenkin kantakaupungin alueella. Suurimpana vaikeutena kevyen liikenteen määrien hyödyntämisessä malleissa on tiedonkeruun vaikeus. Mallissa käytettävien muuttujien arvojen tulee olla määritelty jollaiselle liittymälle mahdollisimman yhdenmukaisella tavalla, jotta aineisto on vertailukelpoista. Tarkkaa laskentatietoa kevyestä liikenteestä ei oletettavasti ole lähitulevaisuudessa käytettävissä, mutta eräs mahdollisuus on kuvata jalankulkijoiden ja pyöräilijöiden lukumäärää diskreetillä asteikolla, esimerkiksi kokonaisluvulla välillä 0-5 tai 0-10. Asteikon lukuarvo voi perustua liikennelaskentaan tai arvioon kevyen liikenteen määrästä liittymässä. Myös tämän yksinkertaistetun lähestymistavan vaikeutena on tietojen kerääminen

yhdenmukaisella tavalla kaikista liittymistä.

Nykyisessä liittymätietokannassa ei ole luotettavaa tietoa väistämisvelvollisuudesta tai sen osoittamisesta liittymissä. Väistämisvelvollisuus vaikuttaa kuljettajien käyttäytymiseen liittymässä, joten tieto kärkikolmiosta tai pakollista pysähtymistä osoittavasta merkistä voisi mahdollisesti parantaa mallien ennustamiskykyä valo-ohjauksettomien liittymien tapauksessa.

Liikennemäärätiedon jatkuva ja automaattinen kerääminen parantaisi mallien luotettavuutta etenkin vähäliikenteisissä liittymissä. Liikennemäärätietojen keräämisen kustannukset eivät ole tässä vaiheessa perusteltuja pelkän onnettomuusmallin vuoksi, mutta tulevaisuudessa liikennemäärätietojen automaattinen ja reaaliaikainen kerääminen voi muuttua nykyistä helpommaksi ja edullisemmaksi. Samoja, aiempaa tarkempia tietoja olisi mahdollista hyödyntää myös muussa suunnittelussa.

Lähteet

- K. Burnham & D. Anderson (2003). *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. Springer. ISBN 9780387953649.
- W. Cheng & S. P. Washington (2005). ‘Experimental evaluation of hotspot identification methods’. *Accident Analysis & Prevention* **37**(5):870–881.
- R. Elvik (2009). *The Handbook of Road Safety Measures*. Emerald Group Publishing. ISBN 9781848552500.
- Euroopan komissio (2010). *Kohti eurooppalaista tieliikenneturvallisuuksaluetta: tieliikenneturvallisuuuden poliittiset suuntaviivat 2011-2020*.
- N. J. Garber & G. Rivera (2010). ‘Safety performance functions for intersection on highways maintained by the Virginia Department of Transportation’ .
- K. Geurts & G. Wets (2003). *Black spot analysis methods: Literature review*. Onderzoekslijn Kennis Verkeersonveiligheid. Policy Research Centre for Traffic Safety.
- D. Gujarati & D. Porter (2009). *Essentials of Econometrics*. McGraw-Hill higher education. McGraw-Hill Education. ISBN 9780073375847.
- E. Hauer (1997). *Observational Before/After Studies in Road Safety. Estimating the Effect of Highway and Traffic Engineering Measures on Road Safety*. Emerald Group Publishing Limited. ISBN 9780080430539.

- E. Hauer, et al. (2004). ‘How best to rank sites with promise’. *Transportation Research Record: Journal of the Transportation Research Board* **1897**(1):48–54.
- J. Heikkinen (2005). ‘Yleistetyt lineaariset mallit’. Luentomoniste, Helsingin Yliopisto. Saatavissa: <http://www.rni.helsinki.fi/jmh/glm05/glm05.pdf>.
- Helsingin kaupunki (2012). *Liikenneonnettomuudet Helsingissä vuonna 2011*. Helsingin kaupunkisuunnitteluviraston liikennesuunnitteluosaston selvityksiä 2012:3.
- J. M. Hilbe (2011). *Negative Binomial Regression*. Cambridge University Press. ISBN 9781139500067.
- R. Kulmala (1995). *Safety at rural three- and four-arm junctions : development and application of accident prediction models*. Valtion teknillinen tutkimuskeskus. ISBN 9513847713.
- R. Kulmala & M. Roine (1990). *Pääteiden tasoliittymien onnettomuusmallit*. Valtion teknillinen tutkimuskeskus.
- A. Kusumawati & Y. D. Wong (2010). ‘The application of Empirical Bayes approach for identifying and ranking hazardous junctions case study: Singapore signalized junctions.’. *Journal of the Eastern Asia Society for Transportation Studies* **8**:2062–2077.
- Liikenne- ja viestintäministeriö (2010). *Tieliikenteen turvallisuus. Liikenneturvallisuussuunnitelman 2011-2014 taustaraportti*. Julkaisuja-sarja 35/2010. ISBN 9789522431790.
- Liikenneturva (2005). *Liikenneonnettomuuksien tilastointi, selvitys nykytilasta ja kehittämistarpeista*. Liikenneturvallisuuden pitkän aikavälin tutkimus- ja kehittämisohjelma LINTU-julkaisuja 8/2005. Saatavissa: <http://www.lintu.info/LONTTI.pdf>.

- Liikenneturva (2012). *Liikenneonnettomuuksien vakavuuden tilastoinnin kehittäminen*. Liikenneturvallisuuuden pitkän aikavälin tutkimus- ja kehittämisohjelma LINTU-julkaisuja 5/2012. Saatavissa: <http://www.lintu.info/KUUVA.pdf>.
- D. Lord & F. Mannering (2010). 'The statistical analysis of crash-frequency data: a review and assessment of methodological alternatives'. *Transportation Research Part A: Policy and Practice* **44**(5):291–305.
- D. Lord, et al. (2007). 'Further notes on the application of zero-inflated models in highway safety'. *Accident Analysis & Prevention* **39**(1):53–57.
- D. Lord, et al. (2005). 'Poisson, Poisson-gamma and zero-inflated regression models of motor vehicle crashes: balancing statistical fit and theory'. *Accident Analysis & Prevention* **37**(1):35–46.
- J.-L. Martin (2002). 'Relationship between crash rate and hourly traffic flow on interurban motorways'. *Accident Analysis & Prevention* **34**(5):619–629.
- S.-P. Miaou & D. Lord (2003). 'Modeling traffic crash-flow relationships for intersections: dispersion parameter, functional form, and Bayes versus empirical Bayes methods'. *Transportation Research Record: Journal of the Transportation Research Board* **1840**(1):31–40.
- A. Montella (2010). 'A comparative analysis of hotspot identification methods'. *Accident Analysis & Prevention* **42**(2):571–581.
- L. Mountain & B. Fawaz (1996). 'Estimating accidents at junctions using routinely-available input data'. *Traffic engineering & control* **37**(11):624–628.
- B. Persaud, et al. (2010). 'Comparison of empirical Bayes and full Bayes approaches for before–after road safety evaluations'. *Accident Analysis & Prevention* **42**(1):38–43.

- M. Salusjärvi (1992). *Liikenneturvallisuuden käsikirja: katsaus 84 liikenneturvallisuustoimenpiteen vaikutuksiin, kustannuksiin ja virallisiin vastuusuhteisiin*. Liikenneministeriö. ISBN 9789514768217.
- H. Tijms (2003). *A First Course in Stochastic Models*. Wiley. ISBN 9780470864289.
- Tilastokeskus (2015). *Tieliikenneonnettomuustilasto [verkkojulkaisu]*. Suomen virallinen tilasto (SVT). Saatavissa: <http://www.stat.fi/til/ton/>.
- W. Venables & B. Ripley (2002). *Modern Applied Statistics with S*. Statistics and Computing. Springer. ISBN 9780387954578.
- J. M. Ver Hoef & P. L. Boveng (2007). 'Quasi-Poisson vs. negative binomial regression: how should we model overdispersed count data?'. *Ecology* **88**(11):2766–2772.

Liite A

Onnettomuusmallit

Tässä liitteessä on esitetty yksityiskohtaiset tulosteet työssä vertailluista onnettomuusmalleista. Taulukoissa on esitetty regressiokertoimien arvot sekä keskihajonnat. Tähdillä merkitään regressiokertoimien merkitsevyystasoa, ja kertoimet joille on merkitty vain yksi tähti tai ei lainkaan tähtiä eivät ole tilastollisesti merkitseviä yleisesti käytetyillä merkitsevyystasoilla.

Taulukko A.1: Kantakaupungin valo-ohjauksettomat liittymät, malli 1

<i>Dependent variable:</i>	
Onnettomuusfrekvenssi	
Paasuunta	0.00005*** (0.00000)
Constant	−0.752*** (0.027)
Observations	5,145
Log Likelihood	−5,606.879
θ	1.093*** (0.071)
Akaike Inf. Crit.	11,217.760
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01

Taulukko A.2: Kantakaupungin valo-ohjauksettomat liittymät, malli 2

<i>Dependent variable:</i>	
	Onnettomuusfrekvenssi
Paasuunta	0.00004*** (0.00000)
Sivusuunta	0.0001*** (0.00001)
Constant	−0.813*** (0.028)
Observations	5,145
Log Likelihood	−5,577.829
θ	1.112*** (0.071)
Akaike Inf. Crit.	11,161.660
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01

Taulukko A.3: Kantakaupungin valo-ohjauksettomat liittymät, malli 3

<i>Dependent variable:</i>	
	Onnettomuusfrekvenssi
Paasuunta	0.00004*** (0.00000)
Sivusuunta	0.0001*** (0.00001)
Nelihaara	0.737*** (0.046)
Constant	−1.101*** (0.034)
Observations	5,145
Log Likelihood	−5,452.852
θ	1.327*** (0.090)
Akaike Inf. Crit.	10,913.700
<i>Note:</i> *p<0.1; **p<0.05; ***p<0.01	

Taulukko A.4: Kantakaupungin valo-ohjauksettomat liittymät, malli 4

<i>Dependent variable:</i>	
	Onnettomuusfrekvenssi
Paasuunta	0.00004*** (0.00000)
Sivusuunta	0.0001*** (0.00001)
Nelihaara	0.737*** (0.046)
NopRaj40	0.097* (0.057)
NopRaj50	−0.108 (0.095)
NopRaj60	−0.740*** (0.176)
NopRaj70	−0.460* (0.263)
NopRaj80	−2.624*** (0.571)
Constant	−1.151*** (0.036)
Observations	5,145
Log Likelihood	−5,429.845
θ	1.386*** (0.097)
Akaike Inf. Crit.	10,877.690
<i>Note:</i> *p<0.1; **p<0.05; ***p<0.01	

Taulukko A.5: Kantakaupungin valo-ohjauksettomat liittymät, malli 5

<i>Dependent variable:</i>	
	Onnettomuusfrekvenssi
Paasuunta	0.0001*** (0.00000)
Sivusuunta	0.0002*** (0.00001)
Nelihaara	0.742*** (0.045)
NopRaj40	−0.095* (0.058)
NopRaj50	−0.430*** (0.094)
NopRaj60	−0.904*** (0.167)
NopRaj70	−0.436* (0.247)
NopRaj80	−0.825 (0.572)
Paasuunta:Sivusuunta	−0.000*** (0.000)
Constant	−1.418*** (0.041)
Observations	5,145
Log Likelihood	−5,319.043
θ	1.684*** (0.130)
Akaike Inf. Crit.	10,658.090

Note:

*p<0.1; **p<0.05; ***p<0.01

Taulukko A.6: Kantakaupungin valo-ohjauksettomat liittymät, malli 6

	<i>Dependent variable:</i>
	Onnettomuusfrekvenssi
Paasuunta	0.0001*** (0.00000)
Sivusuunta	0.0002*** (0.00001)
Nelihaara	0.697*** (0.075)
NopRaj40	-0.083 (0.059)
NopRaj50	-0.413*** (0.096)
NopRaj60	-0.873*** (0.168)
NopRaj70	-0.273 (0.288)
NopRaj80	-0.866 (0.573)
Paasuunta:Sivusuunta	-0.000*** (0.000)
Paasuunta:Nelihaara	-0.00000 (0.00001)
Sivusuunta:Nelihaara	0.00004 (0.00003)
Paasuunta:Sivusuunta:Nelihaara	-0.000 (0.000)
Constant	-1.407*** (0.045)
Observations	5,145
Log Likelihood	-5,318.080
θ	1.680*** (0.129)
Akaike Inf. Crit.	10,662.160

Note:

*p<0.1; **p<0.05; ***p<0.01

Taulukko A.7: Kantakaupungin valo-ohjauksettomat liittymät, malli 7

	<i>Dependent variable:</i>
	Onnettomuusfrekvenssi
Paasuunta	0.0002*** (0.00002)
Sivusuunta	0.0003*** (0.00004)
Nelihaara	−0.056 (0.144)
NopRaj40	0.491*** (0.120)
NopRaj50	0.167 (0.234)
NopRaj60	−0.563 (1.294)
NopRaj70	−93.705*** (15.845)
NopRaj80	8.596*** (2.087)
Paasuunta:Sivusuunta	−0.00000*** (0.000)
Paasuunta:Nelihaara	0.0001*** (0.00004)
Sivusuunta:Nelihaara	0.001*** (0.0001)
Paasuunta:NopRaj40	−0.0001*** (0.00002)
Paasuunta:NopRaj50	−0.0001*** (0.00002)
Paasuunta:NopRaj60	−0.0001*** (0.00004)
Paasuunta:NopRaj70	0.001*** (0.0003)
Observations	5,145
Log Likelihood	−5,198.813
θ	2.280*** (0.217)
Akaike Inf. Crit.	10,463.630

Note:

*p<0.1; **p<0.05; ***p<0.01

Taulukko A.8: Kantakaupungin valo-ohjauksettomat liittymät, malli 7

	<i>Dependent variable:</i>
	Onnettomuusfrekvenssi
Sivusuunta:NopRaj40	−0.0002*** (0.00004)
Sivusuunta:NopRaj50	−0.0001* (0.00005)
Sivusuunta:NopRaj60	−0.00005 (0.0001)
Sivusuunta:NopRaj70	0.007*** (0.001)
Sivusuunta:NopRaj80	
Nelihaara:NopRaj40	0.919*** (0.246)
Nelihaara:NopRaj50	1.634*** (0.614)
Nelihaara:NopRaj60	19.385*** (2.945)
Nelihaara:NopRaj70	
Nelihaara:NopRaj80	
Observations	5,145
Log Likelihood	−5,198.813
θ	2.280*** (0.217)
Akaike Inf. Crit.	10,463.630

Note:

*p<0.1; **p<0.05; ***p<0.01

Taulukko A.9: Kantakaupungin valo-ohjauksettomat liittymät, malli 7

	<i>Dependent variable:</i>
	Onnettomuusfrekvenssi
Sivusuunta:Nelihaara:NopRaj40	−0.001*** (0.0001)
Sivusuunta:Nelihaara:NopRaj50	−0.001*** (0.0002)
Sivusuunta:Nelihaara:NopRaj60	
Sivusuunta:Nelihaara:NopRaj70	
Sivusuunta:Nelihaara:NopRaj80	
Paasuunta:Sivusuunta:Nelihaara:NopRaj40	0.00000*** (0.00000)
Paasuunta:Sivusuunta:Nelihaara:NopRaj50	0.00000*** (0.00000)
Paasuunta:Sivusuunta:Nelihaara:NopRaj60	
Paasuunta:Sivusuunta:Nelihaara:NopRaj70	
Paasuunta:Sivusuunta:Nelihaara:NopRaj80	
Constant	−1.713*** (0.062)
Observations	5,145
Log Likelihood	−5,198.813
θ	2.280*** (0.217)
Akaike Inf. Crit.	10,463.630

Note:

*p<0.1; **p<0.05; ***p<0.01

Taulukko A.10: Kantakaupungin liikennevalo-ohjauksiset liittymät, malli 1

<i>Dependent variable:</i>	
Onnettomuusfrekvenssi	
Paasuunta	0.00001*** (0.00000)
Constant	0.876*** (0.045)
Observations	1,722
Log Likelihood	−3,758.511
θ	2.120*** (0.133)
Akaike Inf. Crit.	7,521.022
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01

Taulukko A.11: Kantakaupungin liikennevalo-ohjauksiset liittymät, malli 2

<i>Dependent variable:</i>	
	Onnettomuusfrekvenssi
Paasuunta	0.00000 (0.00000)
Sivusuunta	0.00002*** (0.00000)
Constant	0.820*** (0.045)
Observations	1,722
Log Likelihood	-3,735.710
θ	2.230*** (0.143)
Akaike Inf. Crit.	7,477.421
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01

Taulukko A.12: Kantakaupungin liikennevalo-ohjauksiset liittymät, malli 3

<i>Dependent variable:</i>	
	Onnettomuusfrekvenssi
Paasuunta	0.00001*** (0.00000)
Sivusuunta	0.00003*** (0.00000)
Nelihaara	0.534*** (0.046)
Constant	0.350*** (0.059)
Observations	1,722
Log Likelihood	-3,667.861
θ	2.648*** (0.185)
Akaike Inf. Crit.	7,343.721
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01

Taulukko A.13: Kantakaupungin liikennevalo-ohjauksiset liittymät, malli 4

<i>Dependent variable:</i>	
	Onnettomuusfrekvenssi
Paasuunta	0.00002*** (0.00000)
Sivusuunta	0.00004*** (0.00000)
Nelihaara	0.604*** (0.045)
NopRaj40	−0.374*** (0.060)
NopRaj50	−0.675*** (0.071)
NopRaj60	−1.042*** (0.114)
NopRaj80	−0.768*** (0.297)
Constant	0.436*** (0.066)
Observations	1,722
Log Likelihood	−3,611.236
θ	3.012*** (0.223)
Akaike Inf. Crit.	7,238.472

Note: *p<0.1; **p<0.05; ***p<0.01

Taulukko A.14: Kantakaupungin liikennevalo-ohjauksiset liittymät, malli 5

	<i>Dependent variable:</i>
	Onnettomuusfrekvenssi
Paasuunta	0.00004*** (0.00000)
Sivusuunta	0.0001*** (0.00001)
Nelihaara	0.583*** (0.045)
NopRaj40	−0.413*** (0.059)
NopRaj50	−0.727*** (0.070)
NopRaj60	−1.121*** (0.113)
NopRaj80	−0.521* (0.293)
Paasuunta:Sivusuunta	−0.000*** (0.000)
Constant	0.076 (0.082)
Observations	1,722
Log Likelihood	−3,586.051
θ	3.226*** (0.248)
Akaike Inf. Crit.	7,190.102

Note:

*p<0.1; **p<0.05; ***p<0.01

Taulukko A.15: Kantakaupungin liikennevalo-ohjauksiset liittymät, malli 6

	<i>Dependent variable:</i>
	Onnettomuusfrekvenssi
Paasuunta	0.00004*** (0.00000)
Sivusuunta	0.0001*** (0.00001)
Nelihaara	0.755*** (0.154)
NopRaj40	−0.374*** (0.059)
NopRaj50	−0.723*** (0.070)
NopRaj60	−1.153*** (0.113)
NopRaj80	−0.395 (0.289)
Paasuunta:Sivusuunta	−0.000*** (0.000)
Paasuunta:Nelihaara	−0.00002*** (0.00001)
Sivusuunta:Nelihaara	−0.00001 (0.00002)
Paasuunta:Sivusuunta:Nelihaara	0.000** (0.000)
Constant	0.074 (0.118)
Observations	1,722
Log Likelihood	−3,564.050
θ	3.466*** (0.278)
Akaike Inf. Crit.	7,152.100

Note:

*p<0.1; **p<0.05; ***p<0.01

Taulukko A.16: Kantakaupungin liikennevalo-ohjauksiset liittymät, malli 7

	<i>Dependent variable:</i>
	Onnettomuusfrekvenssi
Paasuunta	0.0001*** (0.00002)
Sivusuunta	0.0003*** (0.0001)
Nelihaara	1.520*** (0.456)
NopRaj40	−0.608 (0.419)
NopRaj50	−0.688 (0.446)
NopRaj60	0.805 (0.610)
NopRaj80	5.805** (2.564)
Paasuunta:Sivusuunta	−0.000*** (0.000)
Paasuunta:Nelihaara	−0.0001** (0.00004)
Sivusuunta:Nelihaara	−0.0003*** (0.0001)
Paasuunta:NopRaj40	0.00000 (0.00002)
Paasuunta:NopRaj50	−0.00002 (0.00002)
Paasuunta:NopRaj60	−0.0001*** (0.00003)
Observations	1,722
Log Likelihood	−3,508.532
θ	4.077*** (0.357)
Akaike Inf. Crit.	7,083.064

Note:

*p<0.1; **p<0.05; ***p<0.01

Taulukko A.17: Kantakaupungin liikennevalo-ohjauksiset liittymät, malli 7

	<i>Dependent variable:</i>
	Onnettomuusfrekvenssi
Sivusuunta:NopRaj40	−0.0002** (0.0001)
Sivusuunta:NopRaj50	−0.0002*** (0.0001)
Sivusuunta:NopRaj60	−0.0003*** (0.0001)
Sivusuunta:NopRaj80	
Nelihaara:NopRaj40	0.028 (0.538)
Nelihaara:NopRaj50	0.237 (0.590)
Nelihaara:NopRaj60	−0.474 (3.719)
Nelihaara:NopRaj80	
Observations	1,722
Log Likelihood	−3,508.532
θ	4.077*** (0.357)
Akaike Inf. Crit.	7,083.064
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01

Taulukko A.18: Kantakaupungin liikennevalo-ohjauksiset liittymät, malli 7

	<i>Dependent variable:</i>
	Onnettomuusfrekvenssi
Sivusuunta:Nelihaara:NopRaj40	0.0003*** (0.0001)
Sivusuunta:Nelihaara:NopRaj50	0.0002** (0.0001)
Sivusuunta:Nelihaara:NopRaj60	0.0001 (0.001)
Sivusuunta:Nelihaara:NopRaj80	
Paasuunta:Sivusuunta:Nelihaara:NopRaj40	−0.00000*** (0.000)
Paasuunta:Sivusuunta:Nelihaara:NopRaj50	−0.00000*** (0.000)
Paasuunta:Sivusuunta:Nelihaara:NopRaj60	−0.000 (0.00000)
Paasuunta:Sivusuunta:Nelihaara:NopRaj80	
Constant	−0.385 (0.354)
Observations	1,722
Log Likelihood	−3,508.532
θ	4.077*** (0.357)
Akaike Inf. Crit.	7,083.064

Note:

*p<0.1; **p<0.05; ***p<0.01